Naval Research Laboratory

Monterey, CA 93943-5502

NRL/MR/7531--01-7242

Using Legendre Functions for Spatial Covariance Approximation and Investigation of Radial Nonisotrophy for NOGAPS Data

*Richard Franke

Atmospheric Dynamics & Prediction Branch
Marine Meteorology Division

January 2001

*Affiliation, Dept. of Mathematics, Naval Postgraduate School

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. Agency Use Only *(Leave Blank)*.	2. Report Date.	3. Report Type and Dates Covered.
	January 2001	Final

4. Title and Subtitle.
Using Legendre Functions for Spatial Covariance Approximation and Investigation of Radial Nonisotrophy for NOGAPS Data

5. Funding Numbers.
PE 0602435N
PN BE-35-2-19

6. Author(s).
Richard Franke

7. Performing Organization Name(s) and Address(es).
Naval Research Laboratory
Marine Meteorology Division
Monterey, CA 93943-5502

8. Performing Organization Reporting Number.
NRL/MR/7531--01-7242

8. Sponsoring/Monitoring Agency Name(s) and Address(es).
Office of Naval Research
Arlington, VA 22217-5660

10. Sponsoring/Monitoring Agency Report Number.

11. Supplementary Notes.

12a. Distribution /Availability Statement.
Approved for public release; distribution unlimited

12b. Distribution Code.

13. Abstract *(Maximum 200 words)*.

This report gives the results of an investigation into the approximation of a two-month history of pressure height innovation data from NOGAPS. Two ideas were investigated.

The first investigation was prompted by the fact that all positive definite functions on the sphere have an expansion in terms of Legendre functions with positive coefficients. It was found that for all levels the approximation of the covariance data for pressure height innovations by Legendre functions led to positive coefficients for up to 25 terms except at the some low and high levels. The approximation was constrained to be near zero at distances greater than about 0.6 radians. The interlevel data was approximated indirectly using thickness approximations and these approximations also predominantly led to positive coefficients except when some of the high and low levels were involved. Positive aspects of the scheme are that the least squares approximation is linear and that the interlevel approximations are obtained by a transformation on the coefficients of the single level and thickness approximations, without regard to parameters related to correlation distance used by some other methods.

A second investigation attempted to determine the radial variation of the innovation data by binning the covariance data in six different radial bins for each distance increment, or bin. This data was then approximated using a special second order autoregressive (SAR2) function with radially varying parameters (except for the intercept, which must be the same for all directions). The results indicated a significant variation in the parameter related to correlation distance, for certain levels and thicknesses. However, the results also showed evidence of insufficient data in some directions for some levels and thicknesses since the total amount of data was now split in six directions. Certain of the thickness approximations appear to indicate a possible lack of flexibility of the SAR2 approximation near zero distances.

14. Subject Terms.
Covariance Functions; Data Assimilation; NOGAPS

15. Number of Pages.
56

16. Price Code.

17. Security Classification of Report.	18. Security Classification of This Page.	19. Security Classification of Abstract.	20. Limitation of abstract.
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	Same as report

Contents

ACKNOWLEDGEMENT

This research was conducted while the author was working for Naval Research Laboratory/Monterey during FY00, supported under Program Element 0602435N. The author expresses his thanks to Edward Barker, who made the arrangements, Nancy Baker, Roger Daley, and Andrew Van Tuyl, all of whom contributed to the effort by supplying data and many useful comments. Soheila Judy did formatting and editing of this manuscript.

1. Introduction

Plans for using a three dimensional variational algorithm for data assimilation at Fleet Numerical Meteorology and Oceanography Center (see Daley and Barker, 2000) require new representations for the prediction error. Because of ever increasing computational power, it should now be possible to take advantage of larger assimilation regions and nonisotropic representations of the spatial covariances and cross-covariances between prediction variables. This document gives a report on an investigation into some aspects of spatial covariance representation. Two different approaches are reported. Both were investigated using real data, the pressure height innovations from the FNMOC global weather prediction model, NOGAPS. The data is from a two-month period, September and October 1999, at time 00 UTC, covering the United States and part of Canada, including data between latitudes 25° and 65°, and between longitudes 230° and 290°. This data was subject to the quality control procedures used by NOGAPS. In a further attempt to eliminate some of the outlying data, all stations that reported fewer than ten days at any given level were eliminated for that level, and all stations with a standard deviation of the innovations greater than eighty at any given level were eliminated for that level. This resulted in almost no data being eliminated at 700 mb to 50 mb levels. We will address the question of how much this affected the estimates of prediction and observation error in Section 3.

The two approaches reported here were taken from somewhat different points of view. The first method is basically restricted to isotropic approximations using Legendre polynomials. The point being made here is that since all positive definite isotropic functions on the sphere have a certain kind of expansion in Legendre polynomials, use of this approximation guarantees that approximations of positive definite functions will in turn be positive definite. In addition, certain problems associated with corrrelation distance interpolation between standard levels simply go away.

The second approach is to investigate the radial nonisotropy of the data. This was accomplished by establishing bins of radial width 30°, as well as in geodesic distance. The data in the radial bins was then fit using a special second order autoregressive function in distance. The variation of correlation distance with direction is detailed and shown graphically. While the data shows some variation with direction, the

complications of using such an approximation while maintaining the required positive definiteness may be difficult.

2. Legendre Polynomial Approximations

All isotropic positive definite functions on the sphere have a representation in terms of Legendre polynomials (see Schoenberg, 1942). To be specific, if $\{P_m(x): m = 0,1,...\}$ represents the set of Legendre polynomials, then positive definite functions on the unit sphere are characterized by having an expansion of the form $\sum_{n=0}^{\infty} a_n P_n(\cos\theta)$, where $a_n \geq 0$, $n = 0...$ and θ is the geodesic distance between the two points (angle between vectors from the origin to the two points).

Because the vertical correlations between prediction errors are to be computed, it is necessary to estimate the interlevel covariances. There has been some discussion about whether the interlevel covariances can be computed directly by forming the interlevel covariance data, and then fitting it with an appropriate function. Perhaps this is best done using the Lönnberg and Hollingsworth (1986) scheme wherein the thicknesses are fit and the interlevel covariances are then deduced by using the appropriate equation. Letting $I_i(x) = \delta p_i^o(x) - \delta p_i^p(x)$ represent the pressure level innovation at location x for level i, then considering the spatial covariance of the thickness from level i to level j we find

$$\mathrm{cov}(I_i(x) - I_j(x), I_i(y) - I_j(y)) =$$

$$\mathrm{cov}(I_i(x), I_i(y)) + \mathrm{cov}(I_j(x), I_j(y)) - \mathrm{cov}(I_i(x), I_j(y)) - \mathrm{cov}(I_j(x), I_i(y)).$$

Assuming homogeneity, isotropy, and identical but independent observation instruments, the above right side simplifies to

$$\sigma^2_{o,i,i}\delta_{|x-y|} + C_{i,i}(|x-y|) + \sigma^2_{o,j,j}\delta_{|x-y|} + C_{j,j}(|x-y|) - 2\sigma^2_{o,i,j}\delta_{|x-y|} - 2C_{i,j}(|x-y|),$$

where $C_{i,j}(d)$ represents the spatial covariance of the prediction errors between the i^{th} and j^{th} levels at distance $d = |x-y|$, $\sigma^2_{o,i,j}$ represents the covariance of the observation errors between the i^{th} and j^{th} levels, and δ_d is the Kronecker delta.

Solving for the interlevel spatial covariance of the prediction errors, we then have

$$C_{i,j}(d) = 0.5\Big[\sigma^2_{o,i,i}\delta_d + C_{i,i}(d) + \sigma^2_{o,j,j}\delta_d + C_{j,j}(d) - 2\sigma^2_{o,i,j}\delta_d -$$

$$\text{cov}(I_i(x) - I_j(x), I_i(y) - I_j(y))\Big]$$

For nonzero distances the Kronecker delta is zero, so assuming the approximation of only $x \neq y$ data for the last term, we find that

$$C_{i,j}(d) = 0.5\Big[C_{i,i}(d) + C_{j,j}(d) - \text{cov}(I_i(x) - I_j(x), I_i(y) - I_j(y))\Big]$$

We now approximate the single level data using the Legendre approximation, obtaining $C_{i,i}(d) \approx \sum_{n=0}^{N} a_n^{i,i} P_n(\cos\theta)$. The upper limit N is some maximum number of terms, with $a_n^{i,i}$ perhaps being zero beyond some value depending on the level. We then approximate the interlevel thicknesses, obtaining $\text{cov}(I_i(x) - I_j(x), I_i(y) - I_j(y)) \approx \sum_{n=0}^{N} b_n^{i,j} P_n(\cos\theta)$, again with the $b_n^{i,j}$ perhaps being zero beyond some value depending on the two levels. This then leads to the interlevel covariance approximation

$$C_{i,j}(d) \approx \sum_{n=0}^{N} a_n^{i,j} P_n(\cos\theta) = 0.5\left[\sum_{n=0}^{N} a_n^{i,i} P_n(\cos\theta) + \sum_{n=0}^{N} a_n^{j,j} P_n(\cos\theta) - \sum_{n=0}^{N} b_n^{i,j} P_n(\cos\theta)\right] =$$

$$0.5\sum_{n=0}^{N} (a_n^{i,i} + a_n^{j,j} - b_n^{i,j}) P_n(\cos\theta),$$

or $a_n^{i,j} = 0.5(a_n^{i,i} + a_n^{j,j} - b_n^{i,j})$. For zero distance, the empirical vertical covariance matrix for the prediction errors is then given by the matrix $\left[\sum_{n=0}^{N} a_n^{i,j}\right]$.

Because it is necessary that statistical objective analysis schemes have a representation of the spatial covariance of the prediction error, an investigation of the efficacy of using the Legendre representation of the prediction error was undertaken. The data was as described in the Introduction. While the data set itself had distances as large as 0.7+ radians, all data for distances greater than 0.6 radians were discarded. It is almost certainly true that the correlation tends to zero as the distances become large. In addition this is very desirable from a computational point of view. In an attempt to ensure this,

zero valued data at a spacing of 0.05 was added starting at distance 0.64159 out to distance π.

In line with previous studies by the author (Franke, 1999a, 1999b; Franke and Barker, 2000), the binned data was weighted by the number of covariance pairs in the bin. The artificial data was given the maximum of the weights for the actual data. The fitting problem then becomes a linear problem, and least squares was used to determine the coefficients in the Legendre expansion. A maximum of 25 coefficients in the expansion was permitted; that is, a maximum 24^{th} degree polynomial was used. However, if any coefficients were negative, the degree was successively decreased until all coefficients in the resulting fit were nonnegative. An alternative method would be to delete the term with (say) the largest negative coefficient and recompute the fit, continuing until all coefficients are nonnegative. No alternate method was explored, however, and in many instances all 25 coefficients were positive. Most (all but three) of the forty-two exceptions involved the 1000mb, 925mb, 30mb, 20mb, and 10mb levels. In only one instance did the same level data lead to fewer than 25 coefficients being positive, this occurring at the 50mb level.

Some of the 136 sets of the binned data and the resulting fits are given. Shown in Figure 1 are four different single level fits. Notice that the curves tend to have small amplitude variations imposed on the expected decaying form. These appear to be real, based on the binned data values, but whether they are physically real or not is unknown to this investigator.

The standard deviations of the prediction errors, observation errors, and innovations (total error) are shown in Figure 2. These curves are somewhat different for this data from 1999 than that taken in 1996, for which the corresponding plots are shown in Franke (1999b, Figures 16 and 17). The error curves shown in Franke (1999b) show a definite hook to larger values below 850 mb. In Figure 2 no such behavior is noted. Above 300 mb the observation error curve in Figure 2 has a slight decrease before continuing to increase above that, and then similar behavior as Figure 17 (time 00 UTC, as is the time for Figure 2) in Franke (1999b). However, the observation error takes on slightly smaller values in Figure 2, above 300 mb, generally by two to three meters. The prediction error curve in Figure 2 is also some meters smaller than that in Franke (1999b) Figure 17

above 150 mb. Here the curve decreases up to 50 mb before beginning to increase in value. At 50 mb the value of prediction error is approximately 5 meters smaller.

Binned thickness data and the Legendre fits for four thicknesses are shown in Figure 3. These figures exhibit much the same kind of behavior. Finally, the implied interlevel cross-covariance curves are shown in Figure 4. Also shown are the implied binned data values. These values were obtained by applying the same transformation to the binned values (but plotting them at the midpoints of the bins) and will not be exactly the same points as would be obtained by computing the binned interlevel cross-covariances from the raw data. However, differences would probably be minor. Figure 4 shows the same somewhat wiggly behavior, however the 850mb level to the 250mb level cross-covariance data is more scattered and the curve initially increases in value. Given the data, the curve looks reasonable.

Some information about the behavior of the sequence of coefficients can be seen from the bar charts for the sequences for the prediction errors at the various levels, shown in Figure 5. All levels are plotted with the same vertical scale except the 10mb level. It is interesting to note that there is a bimodal distribution of coefficients for most levels. The 10mb level distribution looks somewhat anomalous compared to the other fifteen levels.

The coefficients of the interlevel cross-covariance functions almost all include negative values; to be specific, 90 out of 120 interlevel cross-covariances included negative coefficients. This indicates that fitting the interlevel cross-covariances with a positive definite function is probably improper.

The vertical correlation of the pressure level prediction error derived from the above fitting procedure is shown in Figure 6, indicated by the marked points. Also shown is the fit obtained by simultaneously finding a transformation of logP and a fit to the data using a full third order autoregressive function in the transformed coordinates. See Sampson and Guttorp (1992) and Franke (1999a, 1999b) for details of this procedure. The fitting values are plotted without symbols using the same type of line. Note that while the vertical correlation matrix derived from the Legendre fits is not positive definite, the fitting function for the vertical correlation data is positive definite. The curves were plotted using piecewise linear interpolation for the transformation curve, hence making the curve as plotted nondifferentiable. However the fitting curve itself, shown in Figure 7

has a very sharp transition at zero distance, yielding very large second derivatives. The sharp peaks shown in the correlation curves may not be entirely attributable to the piecewise linear transformation curve. Figure 7 also shows the transformed data along with the fitting function, while Figure 8 shows the transformation of the logP coordinate system for the prediction errors.

Figures 9, 10, and 11 show the corresponding graphs for the observation errors. As was noted in Franke (1999b), the vertical correlation curves are much broader for the observation error than for the prediction error.

After the vertical correlation matrix for the height errors has been approximated, a new interlevel thickness error fit is implied, one having the appropriate intercept as indicated by the variances of the errors at the two levels and the correlation between the two levels. That is, the empirical covariance matrix is given by $\left[\sum_{n=0}^{N} a_n^{i,j}\right]$, but we have approximated this matrix by $\left[\sum_{n=0}^{N} \alpha_n^{i,j}\right]$. To be consistent, we must now redefine the interlevel fits with the constraint implied by that approximation. Since $\sum_{n=0}^{N} a_n^{i,j} = .5\sum_{n=0}^{N}(a_n^{i,i} + a_n^{j,j} - b_n^{i,j})$, we now redefine the coefficients in the thickness fits (call the new values $\beta_n^{i,j}$) so that $\sum_{n=0}^{N} \alpha_n^{i,j} = .5\sum_{n=0}^{N}(a_n^{i,i} + a_n^{j,j} - \beta_n^{i,j})$. This is achieved by fitting the thickness data subject to the constraint that for each pair of levels, $\sum_{n=0}^{N} \beta_n^{i,j} = \sum_{n=0}^{N}(a_n^{i,i} + a_n^{j,j} - 2\alpha_n^{i,j})$. This, in turn, implies a new interlevel spatial covariance function that satisfies the approximate vertical correlation properties given by Figure 6. The new curves (and the binned data) are shown in Figure 12 for the same levels as shown in Figure 4, to which the curves can be compared. For three of the pairs of levels, the differences are minor. For the 850mb and 250mb levels the curves are quite different from each other, the intercept shown in Figure 12 being forced to be considerably larger than that in Figure 4. Beyond a distance of about 0.175 radians the curves take on essentially similar values.

The advantages and disadvantages of using a Legendre representation for spatial covariances and cross-covariances are worth noting. First, the evaluation of the Legendre polynomials is probably somewhat more expensive than other approximation functions with fewer parameters. The cosine of the radial distance is needed, but this relieves one of computing the actual radial distance, saving the computation of the inverse cosine. Evaluation of the Legendre polynomials can be accomplished using a three term recurrence relation that involves three multiplications (and a subtraction) for each degree larger than two. Hence, evaluation of all polynomials up to 24^{th} degree would require sixty-six multiplications, then followed by the dot product of the coefficient vector with the polynomial vector, of length twenty-five. This assumes the coefficients in the recursion relation are precomputed and stored. It would seem to be an advantage that one could interpolate between levels by interpolating the coefficients between levels. No experiments were carried out to try to determine whether the overall interpolation matrix would be positive definite (experiments similar to those carried out in Franke (1999b)).

3. Direction Dependent Covariance Functions

The directional dependence of the spatial covariance function for pressure height prediction error was investigated. In order to do this the data was collected in bins dependent on both geodesic distance and direction. Covariance value is the same independent of whether one takes direction from the first point to the second, or the second point to the first. Because the direction of shortest distance between two points is the not the opposite for the two points, it is necessary to do an average of directions between the two points. For these experiments it was done as follows. Let the direction from point one to point two be initially on a heading of α degrees from north (positive to the east) and the direction from point two to point one be initially on a heading of β degrees ($-180 < \alpha, \beta \leq 180$). Then the angle of direction between the two points was taken to be $\frac{1}{2}(\alpha + \beta + 180)$. The covariance data for each level was accumulated in 80 distance bins of size 0.01 radian and into radial bins of size 30°.

The spatial covariance was modeled using the special second order autoregressive function, $F(d; A, b) = A(1 + bd)e^{-bd}$. The value of b was allowed to vary with direction,

while the value of A (the intercept) was constrained to be the same in all directions. Given that there were six radial bins this resulted in a fit at each level having seven parameters. Because the available data is now being put into six times as many bins, the percentage of the amount of data for each of the six fits ranges from about 13% to about 22 % of the total, depending on the direction and the level. At the 1000mb level there was a definite lack of data indicated by the binned data plots. At the 50mb and above level the binned data showed considerable scatter, also partly because of a lack of data.

For each of the sixteen standard levels, the radial data was fit as described in the previous paragraph was computed using least squares, weighted by the number of covariance pairs in each bin corresponding to a positive distance. The spatial covariance fits in the six directions are shown for four different pressure levels in Figures 13-16. The intercept determined the prediction variance, σ_k^2 (for the k^{th} level), with the difference between the variance of the innovations and the prediction variance being the observation variance. The curves in these four figures can be compared qualitatively with the four Legendre fit curves shown in Figure 1. The curves are similar at lower levels and vary more at upper levels, although the Legendre fits have more small scale variability.

The standard deviations of the three quantities, as derived from these approximations, as shown as the solid piecewise linear curves in Figure 17. These curves can be compared with those in Figure 2. Doing so, one sees that the prediction error standard deviation curve is very similar in the two cases. At the lower and upper levels the prediction standard deviation is somewhat larger in Figure 17, while the prediction error is somewhat smaller at the intermediate levels. As noted in the introduction, stations reporting fewer than ten days at any given level were eliminated for that level, as well as levels for which the standard deviation of the innovations was greater than eighty. Also shown in Figure 17 as dashed curves are the standard deviations when all valid data is included. At the intermediate levels there is generally no difference. At the lower levels there is a no perceptible difference in the prediction error standard deviation, while the observation error standard deviation varies by up to about 0.9m. At the upper three levels the prediction error varies up to about 1.1m, while the observation error varies by more

than 5m at 30mb and 20mb. Apparently discarding the high variance innovations primarily affects the observation error calculation, not a surprising conclusion.

The thicknesses were fit for all pairs of levels. The results show that at thicknesses involving the 1000mb and 50mb and above levels there is a great deal of scatter, as noted previously. Figures 18-21 show the thickness plots for the same thickness data shown in Figure 3 for the Legendre fits. As for the single level fits, the radially varying data show the scattering effect of fewer data in each direction. The variance $\sigma_{l,k}$, of the k to l level thicknesses tend to be considerably larger using the present scheme compared to those derived from the Legendre fits. In fact, it appears that the present scheme may give more reasonable results than the Legendre scheme, although this thought will be revisited in the next paragraph.

Once the single level approximations and the thickness approximations are computed the implied interlevel spatial covariance functions can be computed by $C_{l,k}(d) = \frac{1}{2}\left(\sigma_l^2(1+b_l d)e^{-b_l d} + \sigma_k^2(1+b_k d)e^{-b_k d} - \sigma_{l,k}^2(1+b_{l,k}d)e^{-b_{k,l}d}\right)$. The plots for the six directions for each of the interlevel spatial covariances corresponding to the thickness plots alluded to before are shown in Figures 22-25. Note that the behavior of the spatial covariance functions seem to be anomalous near zero distances. This seems to be caused by the rather large values of $b_{l,k}$ (more will be said later) obtained in the thickness fits, coupled with the small cross-covariance of the interlevel fits (compared with the Legendre fits shown in Figure 4). This smaller value of the cross-covariance is directly attributable to the larger value of the thickness variance obtained using the SAR2 fits. Thus, while the variance of the thickness fits obtained using SAR2 approximations appear to be more reasonable based on the plots of the approximations and the data, in fact the smaller cross-covariances of the Legendre fits leads to what seem to be more reasonable interlevel approximations.

The reciprocal of the parameter b is referred to here as the "correlation distance", and is the distance at which the argument of the exponential has value minus one, and hence the value of the correlation function at that point is $2e^{-1} \approx .74$. The variation of the correlation distance with direction is shown in Figure 26. The value in the six directions computed (and reflected) are shown connected by solid lines. The dashed lines were

obtained by fitting this data with a function of the form $D_l(\theta) = \frac{a_l}{(1+b_l \cos^2(\theta-\theta_l))}$.

Here the l subscript denotes the level. The initial inclination was to fit an ellipse to the data, however, it is important to be able to evaluate the correlation distance easily for a given angle. Thus the above choice, which looks quite elliptic for moderate values of b_l, was chosen. However, it may not be a good choice if the ratio of correlation distances vary by more than about a factor of about two (corresponding to a value of $b_l = 1$), since a slight pinched shaped occurs then (see the 1000mb level curve in Figure 26).

There is a considerable variation in the shape of the correlation distance curves and the orientation of the major axis. The behavior at the 1000mb level is anomalous. At levels from 925mb to 100mb the longer axis of the oval curves are generally in a direction somewhat west of north, although at 250mb the longer axis is almost north. This implies that for these levels the errors have greater spatial correlation in the north-south direction than in the east-west direction. It seems reasonable that errors would be less highly correlated in the direction of air movement since that brings rapid change, while in the perpendicular direction basically the same kind of conditions propagate in the direction of the movement. At the 70mb to 20mb levels the correlation distance curves are close to circular, the fitted curves more so than the derived data. At 10mb the correlation curve is similar to that at the lower levels. The correlation distance plots for these levels show a great variation in size and shape that is probably not indicative of the actual properties of the prediction error.

The correlation distance plots for all 120 interlevel thicknesses are shown in the thumbnail plots in Figure 27a,b. It is interesting that the correlation distances for many of the thicknesses are small, especially compared to the correlation distances for the single level data. This was noted in passing (without showing the data) when discussing the interlevel cross-covariance curves, but in terms of the coefficients $b_{l,k}$ being larger for the thicknesses. This contributes to the thickness part of the interlevel approximation decaying away rather rapidly. Since this part enters negatively, the sum of the three terms yields the kind of initially increasing curve seen in Figures 22-25.

The vertical correlation matrix for pressure height errors implied by the single level and thickness spatial covariance approximations is illustrated in Figure 28, with the same

format as that for Figure 6. The empirical vertical correlation values are indicated by the marked points. Also shown is the fit obtained by simultaneously finding a transformation of logP and a fit to the data using a full third order autoregressive function in the transformed coordinates. As before, the fitting values are plotted without symbols using the same type of line. Again, the vertical correlation matrix derived from the Legendre fits is not positive definite, but the fitting function for the vertical correlation data is positive definite. The curves are similar to, but not the same, as those of Figure 6. The fitting curve and the transformed data are shown in Figure 29, with the transformation shown in Figure 30. The corresponding information for observation errors is given in Figures 31-33, and these are also seen to be similar to the corresponding Figures for the Legendre fits, given in Figures 9-11.

As with the Legendre approximations, we now show the implied interlevel cross-covariance approximations obtained by replacing the thickness variance obtained from the fitting process with that implied by the vertical correlation function. Unlike with the Legendre approximations, we have only replaced the variance of the thickness, and have not then refit the data with the variance fixed. That is, at zero distance we have $C_{l,k} = \frac{1}{2}\left(\sigma_l^2 + \sigma_k^2 - \sigma_{l,k}^2\right)$. In cases where $C_{l,k}$ is made larger by the vertical correlation approximation (for example, for the 850mb and 250mb levels), a smaller approximate value of $\sigma_{l,k}^2$ is implied, about 84 rather than 100. If a new SAR2 approximation to the data shown in Figure 19 were computed with the intercept fixed at the smaller value, it is certain that the correlation distance would be larger. This in turn would alleviate the problem shown in Figure 35, where the interlevel spatial cross-covariance initially rises sharply from its value at zero distance. On the other hand, the 850mb to 500mb levels spatial cross-covariance would still be essentially like that in Figure 34, which is also essentially like that in Figure 22. This follows since the approximate value of the correlation shown in Figure 28 is essentially the same as the empirical value computed from the thickness variances. The 500mb to 250mb spatial cross-covariance and the 500mb to150mb spatial cross-covariance changes would be similar to the 850mb to 500mb and 850mb to 250mb changes, respectively. The data and implied

approximations after the vertical correlation approximation are shown for the four pairs of levels in Figures 34-37.

4. Conclusions

The results of this study have given rise to some interesting questions. We will emphasize the two approaches separately, while comparing them.

The fact that 25 coefficients of the Legendre approximation routinely (but not always) turn out to be positive indicates that, for the most part, the data is sufficient to yield actual positive definite spatial approximations. The Legendre approximations also confirm that the interlevel spatial cross-covariance functions are not necessarily positive definite, and probably should not be approximated by positive definite functions. The interlevel cross-covariance functions could certainly be approximated using the Legendre functions, however the implied thickness covariance functions would not be constrained to be positive definite. The Hollingsworth and Lönnberg scheme of obtaining implied interlevel approximations via the thickness approximations seems to this investigator to the best idea. Using the approximate vertical correlation functions to yield the thickness variances, followed by the recomputation of the thickness approximation subject to the appropriate variance seems to yield a consistent approximation. However, as noted above, it is unknown whether the full 3D approximation obtained in this manner is positive definite, or not.

One drawback of the approximation is that the evaluation of the approximation may be more time consuming than using other approximations, such as the SAR2 approximation. If an overall more consistent approximation is obtained the additional expense may be worthwhile. The radially varying approximation investigated here using the SAR2 spatial approximation could be carried out using the Legendre approximation. The interpolation between coefficients in the radial direction, as well as the vertical is easily accomplished with the Legendre approximation.

The radially varying spatial covariance aproximations show that there is a pronounced difference in different directions, at least at certain levels. A possible problem is that in order to have sufficient data to well-define the data, a time period long enough that temporal effects are introduced may be necessary. While the two month

results we have seem reasonable, a longer time period may yield better results. If fewer directions were computed, perhaps 4 instead of the 6 used here, the binned data would be somewhat less scattered in most instances.

References

Roger Daley and Edward Barker (2000): The NAVDAS Source Book, NRL Publication NRL/PU/7530--00-418, August 2000.

Hollingsworth and P. Lönnberg (1986): The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field, *Tellus* 38A, 111-136.

Richard Franke (1999a): Vertical correlation functions for temperature and relative humidity errors, Naval Research Laboratory/Monterey TR#NRL/MR/7531-99-7240, January 1999, 69 pp. (NTIS ADA 361021)

Richard Franke (1999b): Three dimensional covariance functions for NOGAPS data, *Mon. Weather Rev.* 127, 2293-2308.

Richard Franke and Edward Barker, 2000: Vertical Correlation functions for temperature and relative humidity errors, *Monthly Weather Review* 128, 3692-3981.

Paul D. Sampson and Peter Guttorp (1992): Nonparametric estimation of nonstationary spatial covariance structure, *J. American Statistical Society* 87, 108-119.

I. J. Schoenberg (1942): Positive definite functions on spheres, *Duke Math. J.* 9, 96-108.

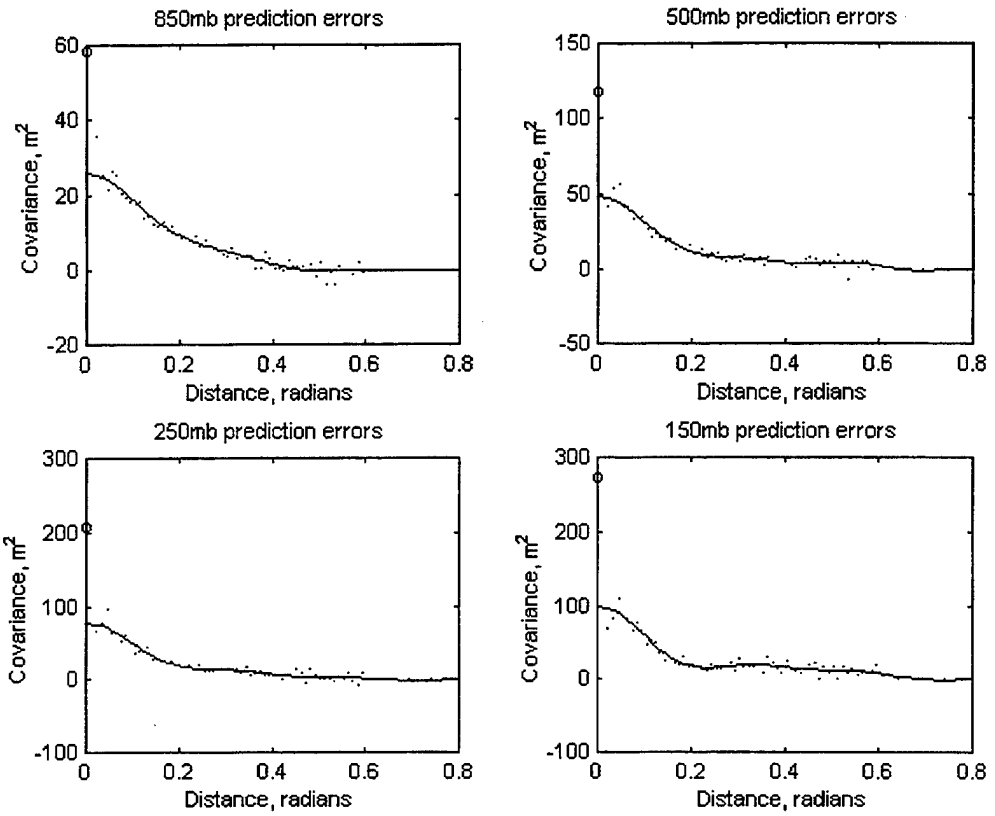

Figure 1: Binned pressure level prediction error at four levels and the fit by Legendre polynomials. The total innovation variance is shown by the "o" symbol on the axis.

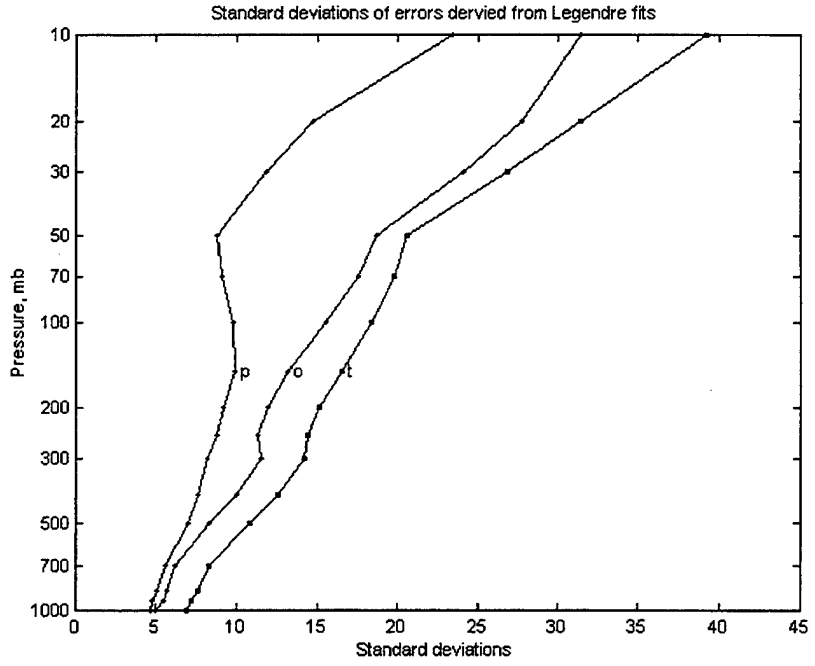

Figure 2: Standard deviations of the prediction error (p), the observation error (o), and the total error (t).

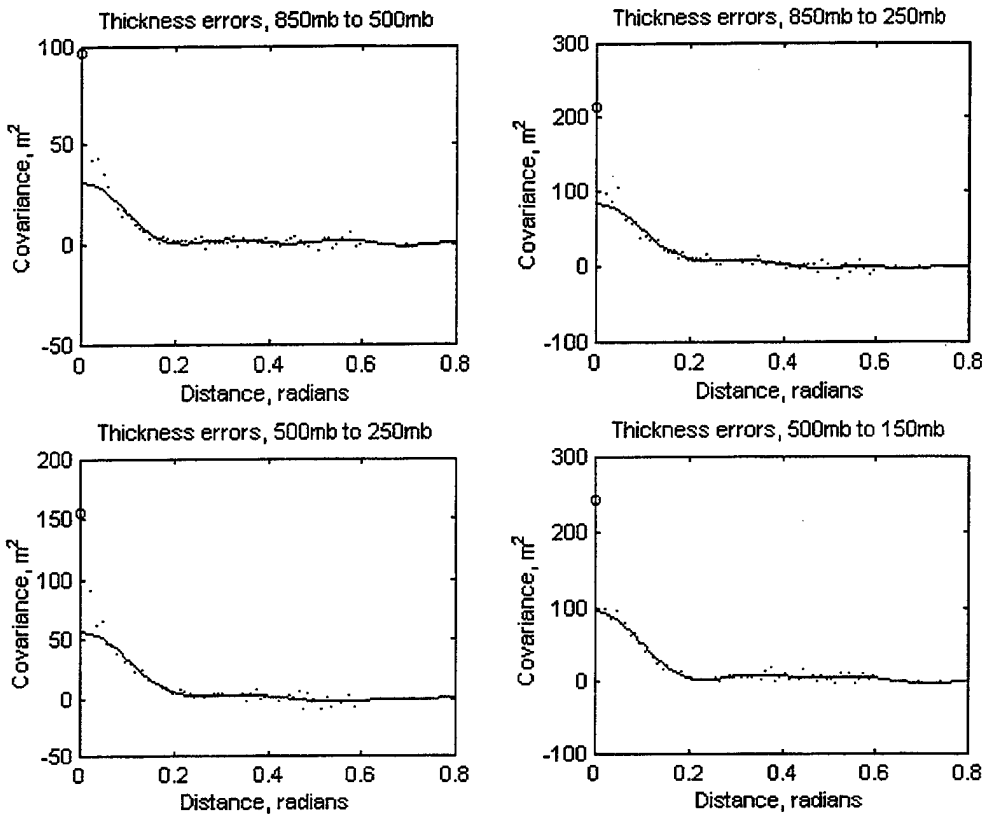

Figure 3: Binned interlevel thickness error data at four pairs of levels and fits by Legendre polynomials. The total innovation variance is shown by the "o" symbol on the axis.

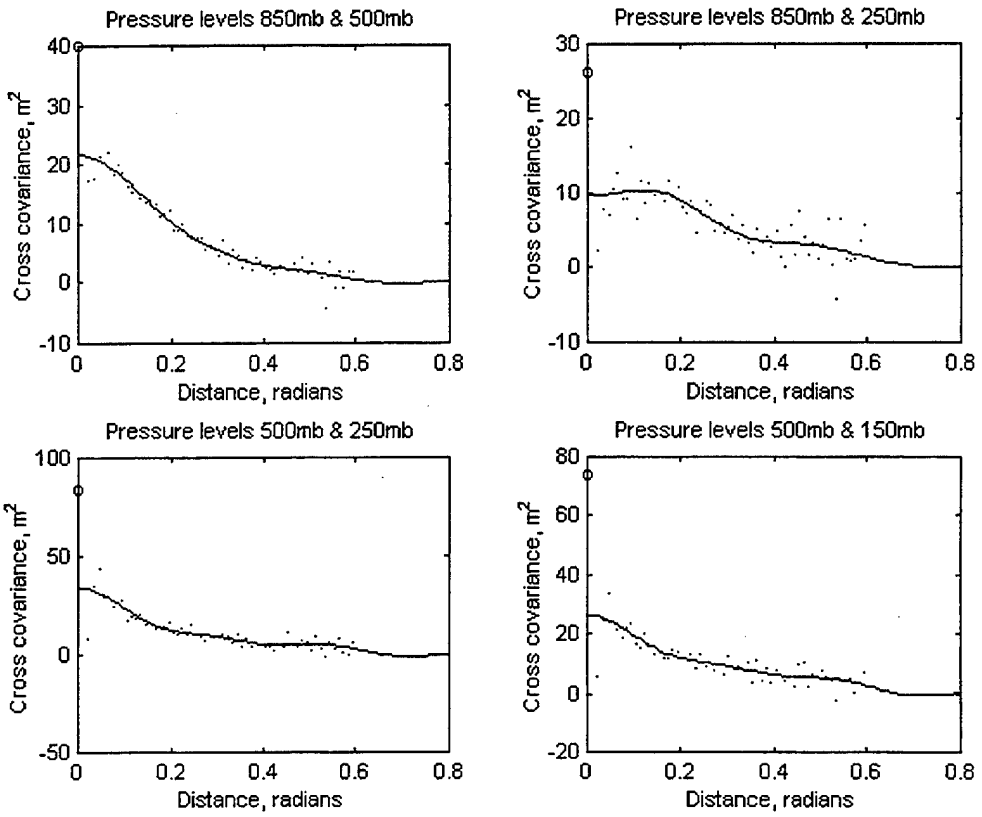

Figure 4: Interlevel pressure level error cross-covariance data from binned single level and binned interlevel thickness error data with implied fit from single level error fits and interlevel thickness error fits using Legendre polynomials. The total innovation cross-covariance is shown by the "o" symbol on the axis.

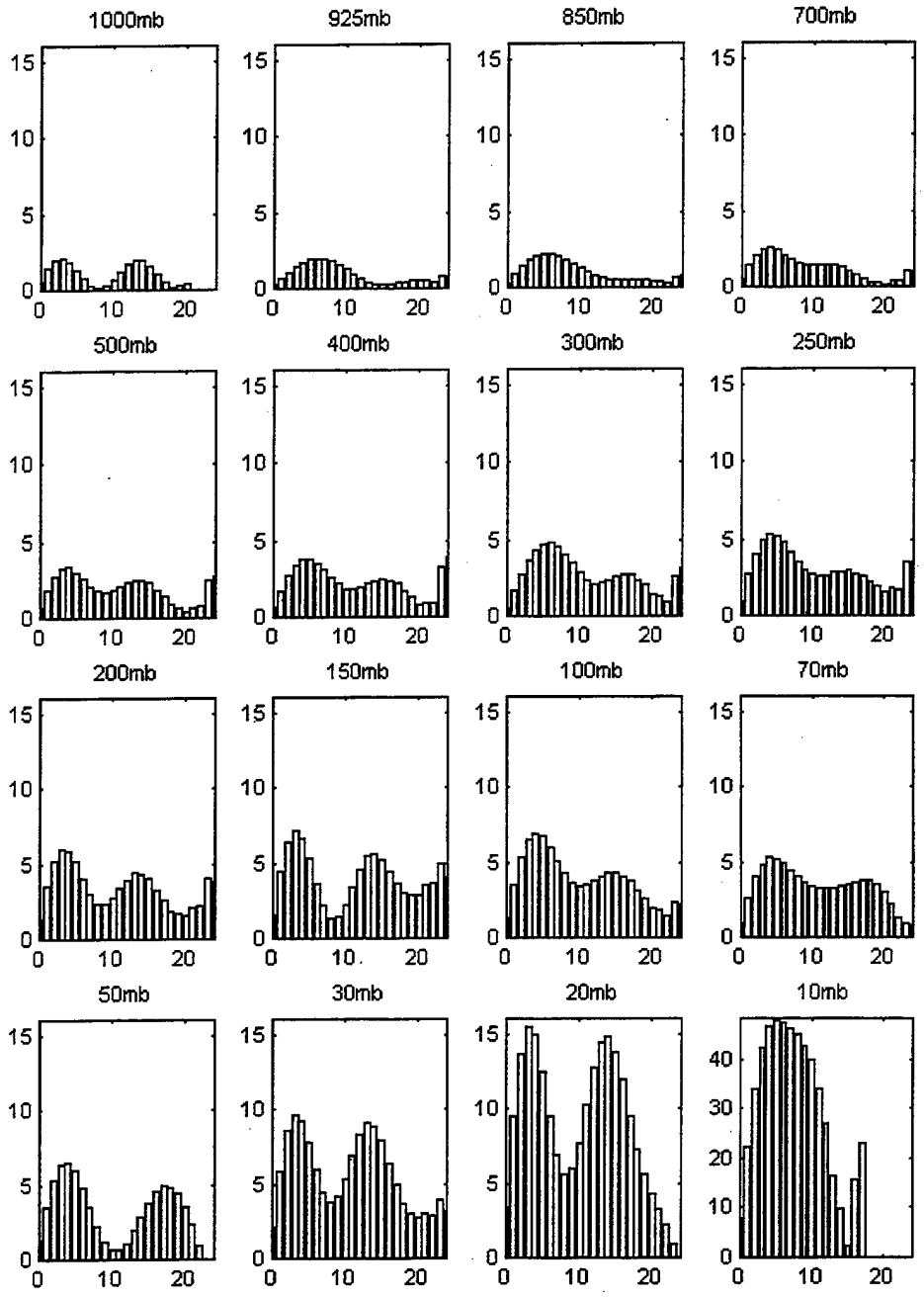

Figure 5: Bar chart showing the values of the a_n coefficients for the sixteen standard levels. All levels have the same vertical scale except for the 10mb level plot.

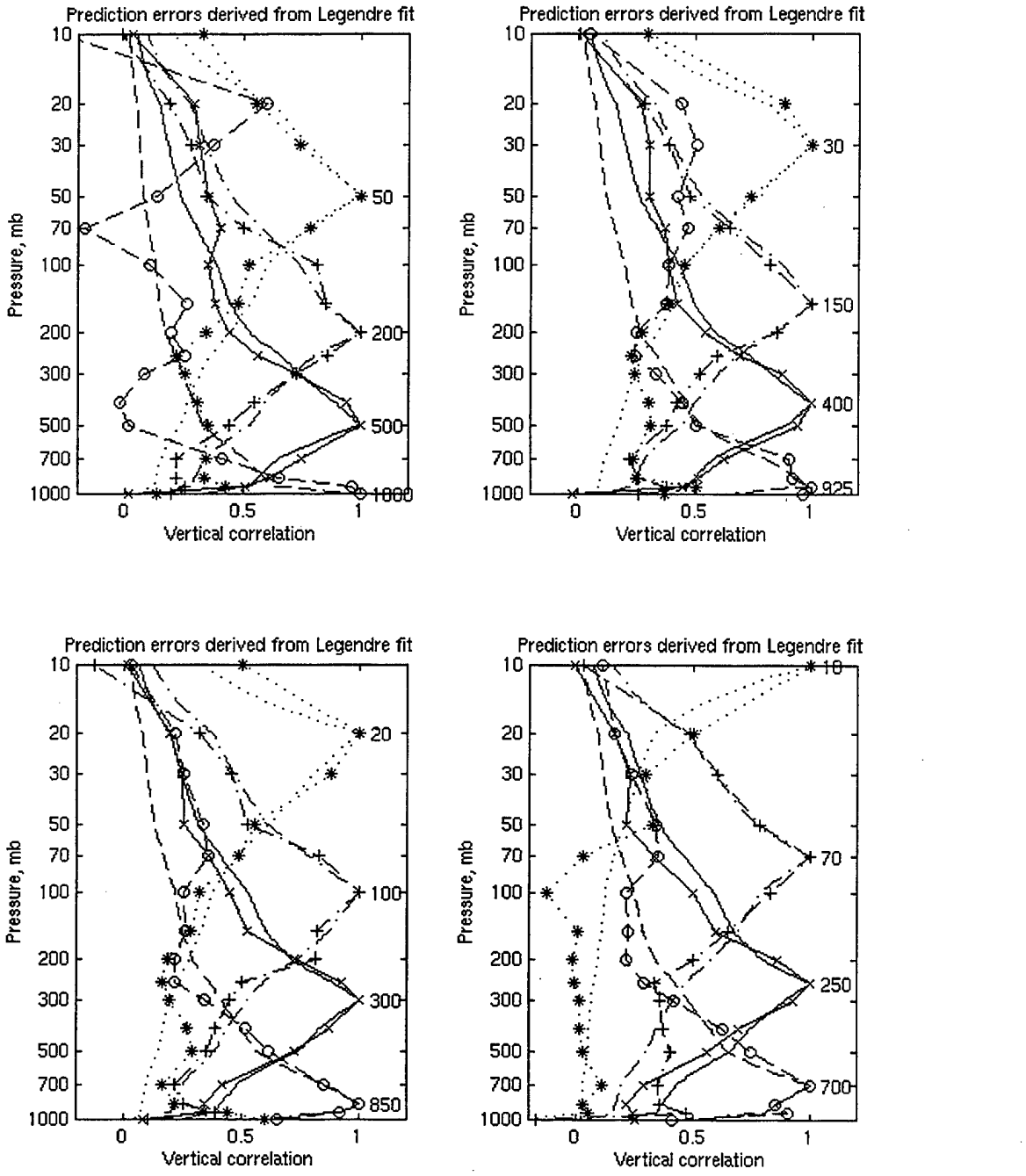

Figure 6: Vertical correlation curves for prediction error obtained from Legendre fits (with symbols) and the approximation obtained by simultaneously finding a transformation of logP and fitting with a full third-order autoregressive function (same type curve, without symbols).

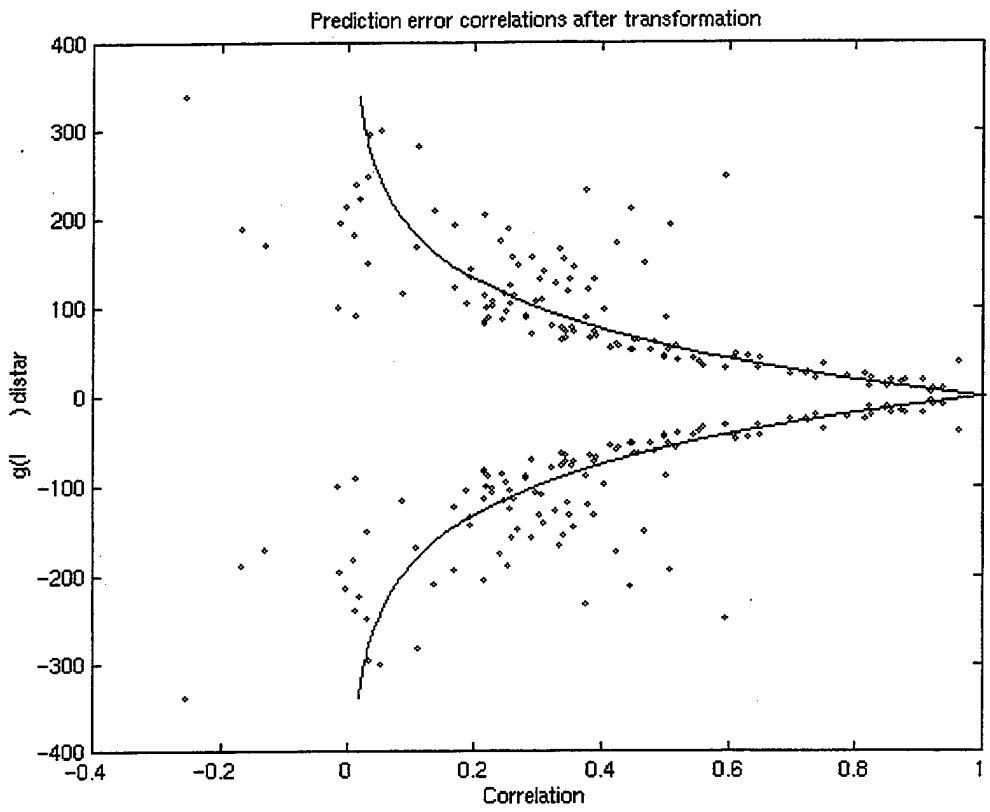

Figure 7: Vertical prediction error correlations and fitting curve for prediction error in the transformed coordinate system.

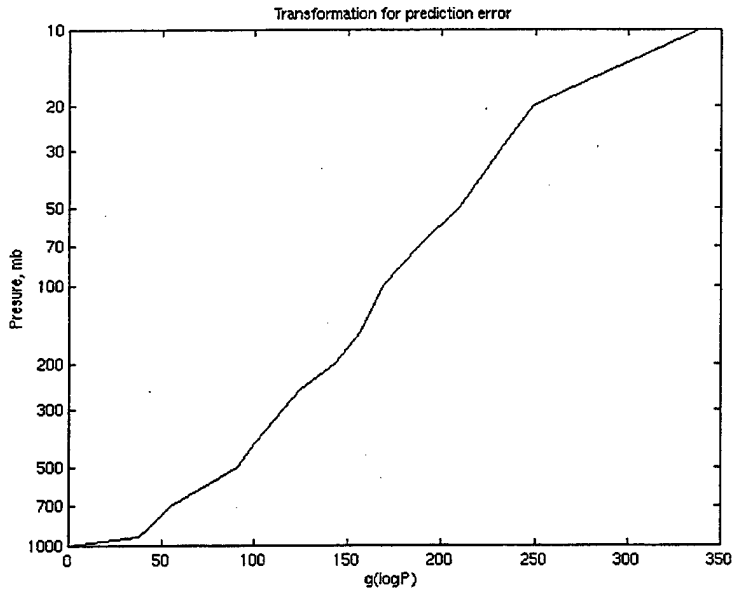

Figure 8: Transformation curve for prediction error.

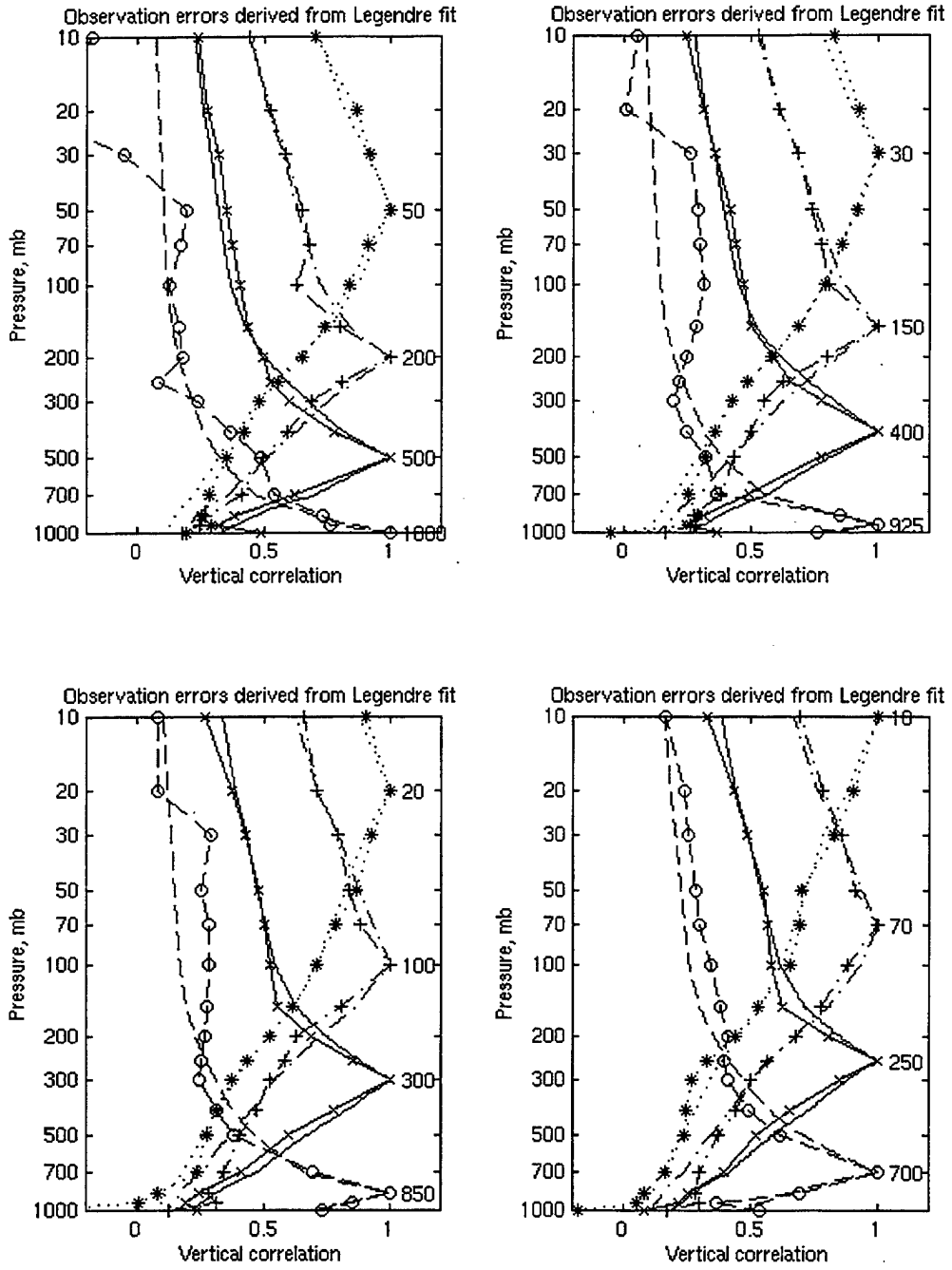

Figure 9: Vertical correlation curves for observation error obtained from Legendre fits (with symbols) and the approximation obtained by simultaneously finding a transformation of logP and fitting with a full third-order autoregressive function (same type curve, without symbols).

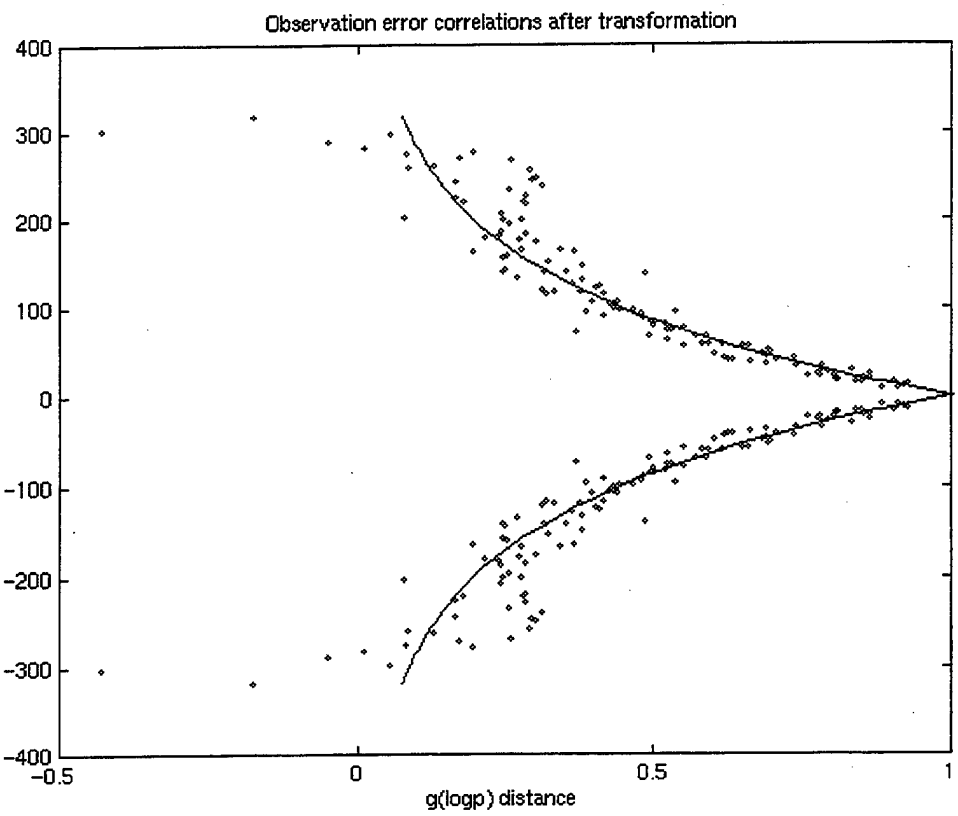

Figure 10: Vertical observation error correlations and fitting curve for observation error in the transformed coordinate system.

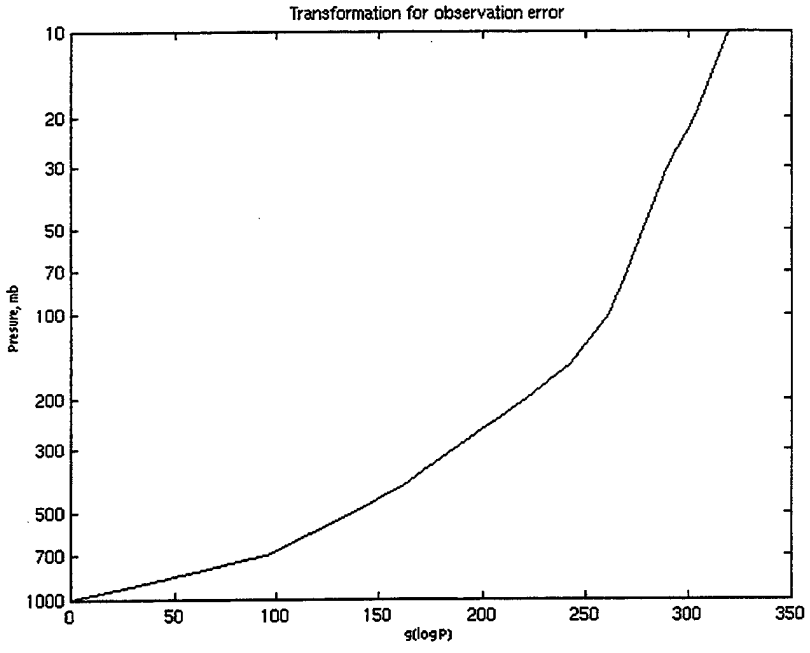

Figure 11: Transformation curve for observation error.

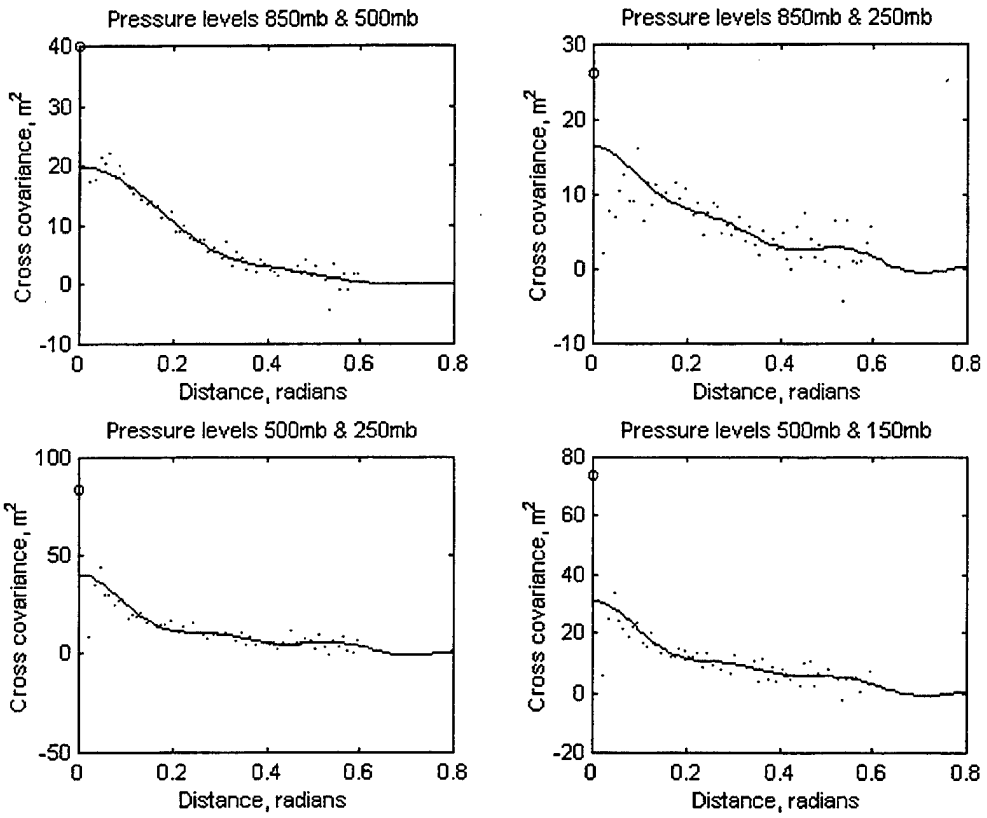

Figure 12: Interlevel pressure level error cross-covariance data from binned single level and binned interlevel thickness error data with implied fit from single level error fits and interlevel thickness error fits using Legendre polynomials after imposing the implied thickness variance derived from the vertical correlation approximations shown in Figure 6. Compare with Figure 4.

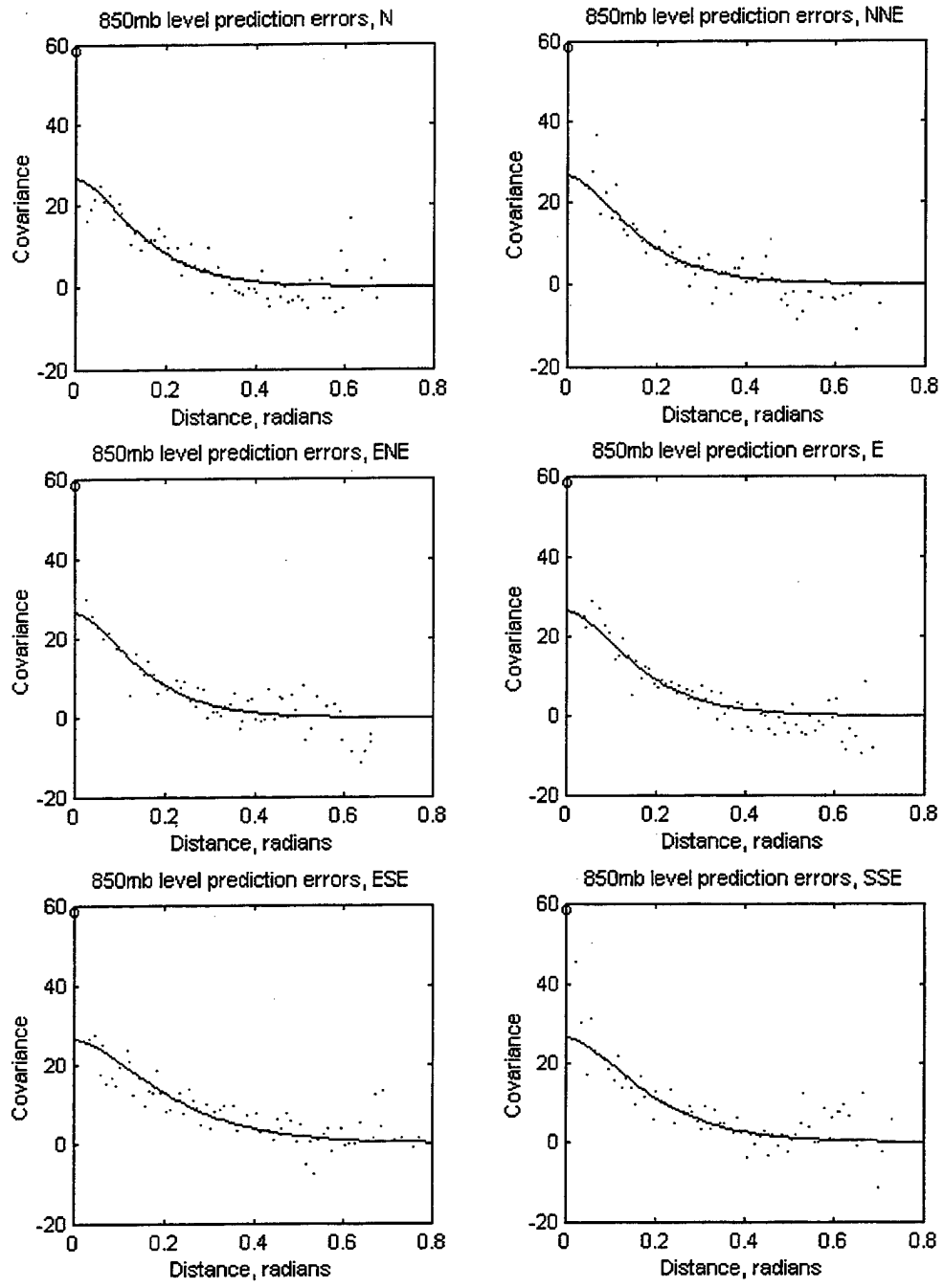

Figure 13: Binned pressure level error covariance data and spatial covariance approximations using SAR2 approximation, in six directions at 30° intervals, starting at North and going clockwise. This data is for the 850 mb level.

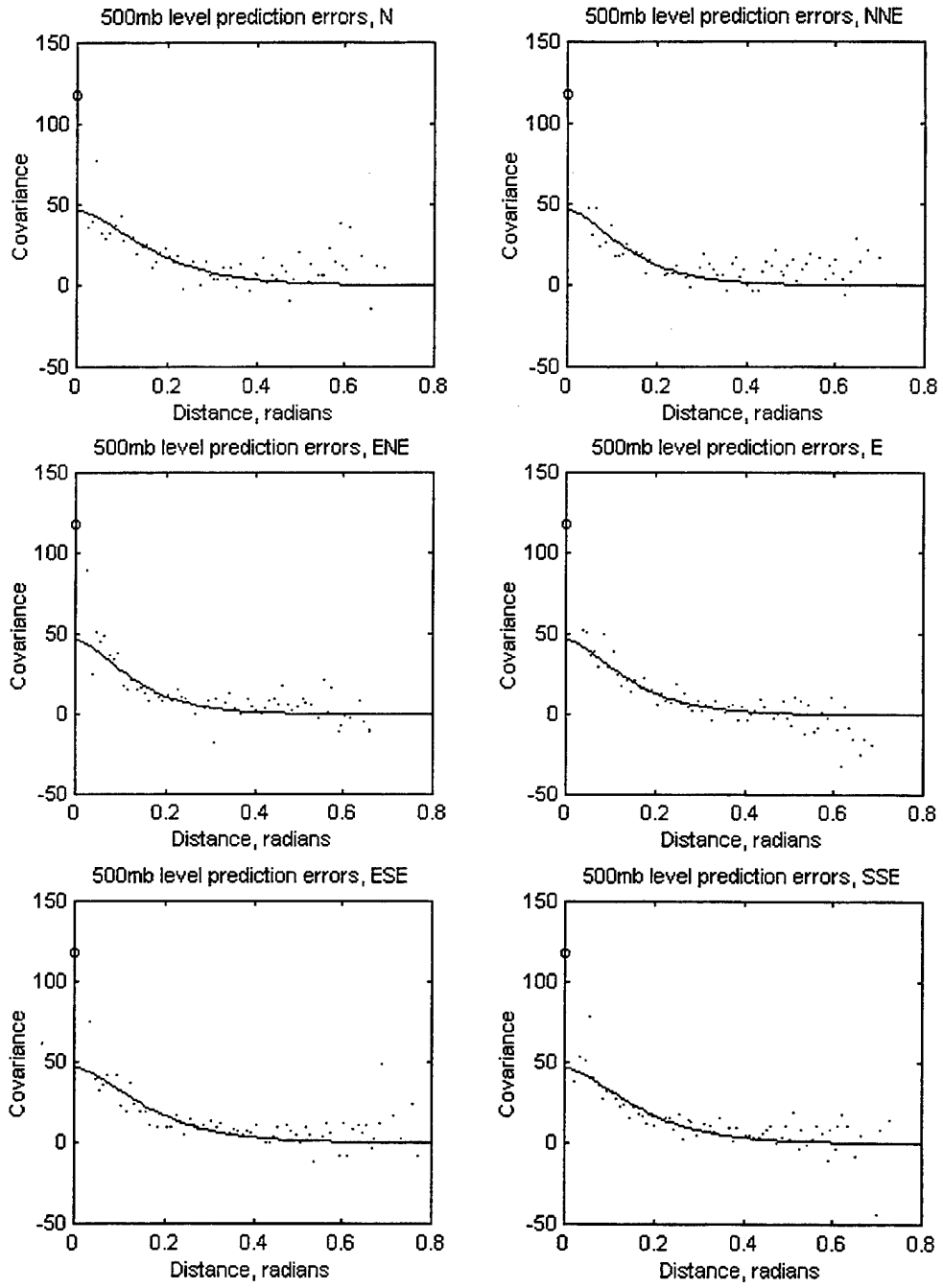

Figure 14: Binned pressure level error covariance data and spatial covariance approximations using SAR2 approximation, in six directions at 30° intervals, starting at North and going clockwise. This data is for the 500 mb level.

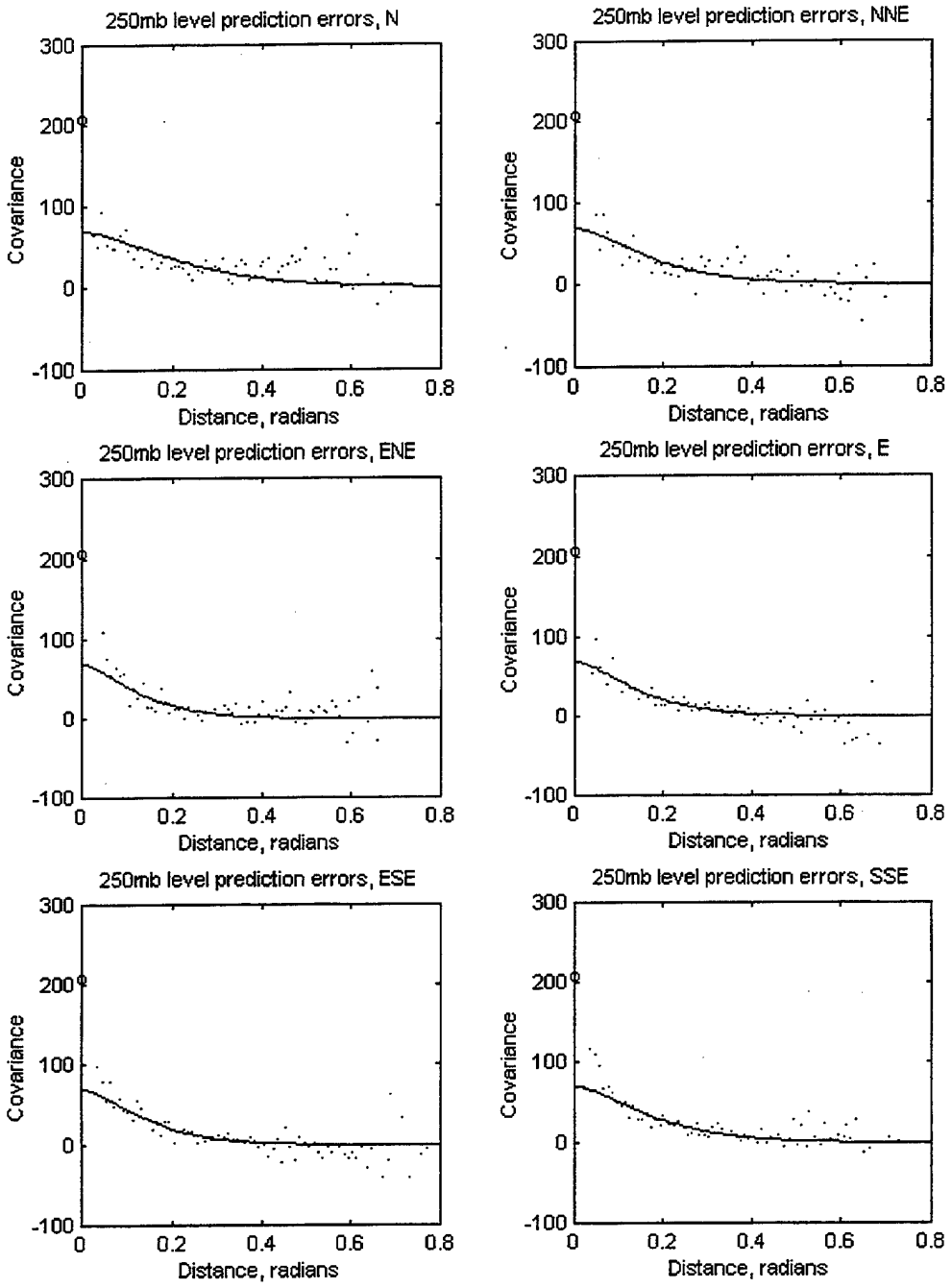

Figure 15: Binned pressure level error covariance data and spatial covariance approximations using SAR2 approximation, in six directions at 30° intervals, starting at North and going clockwise. This data is for the 250 mb level.

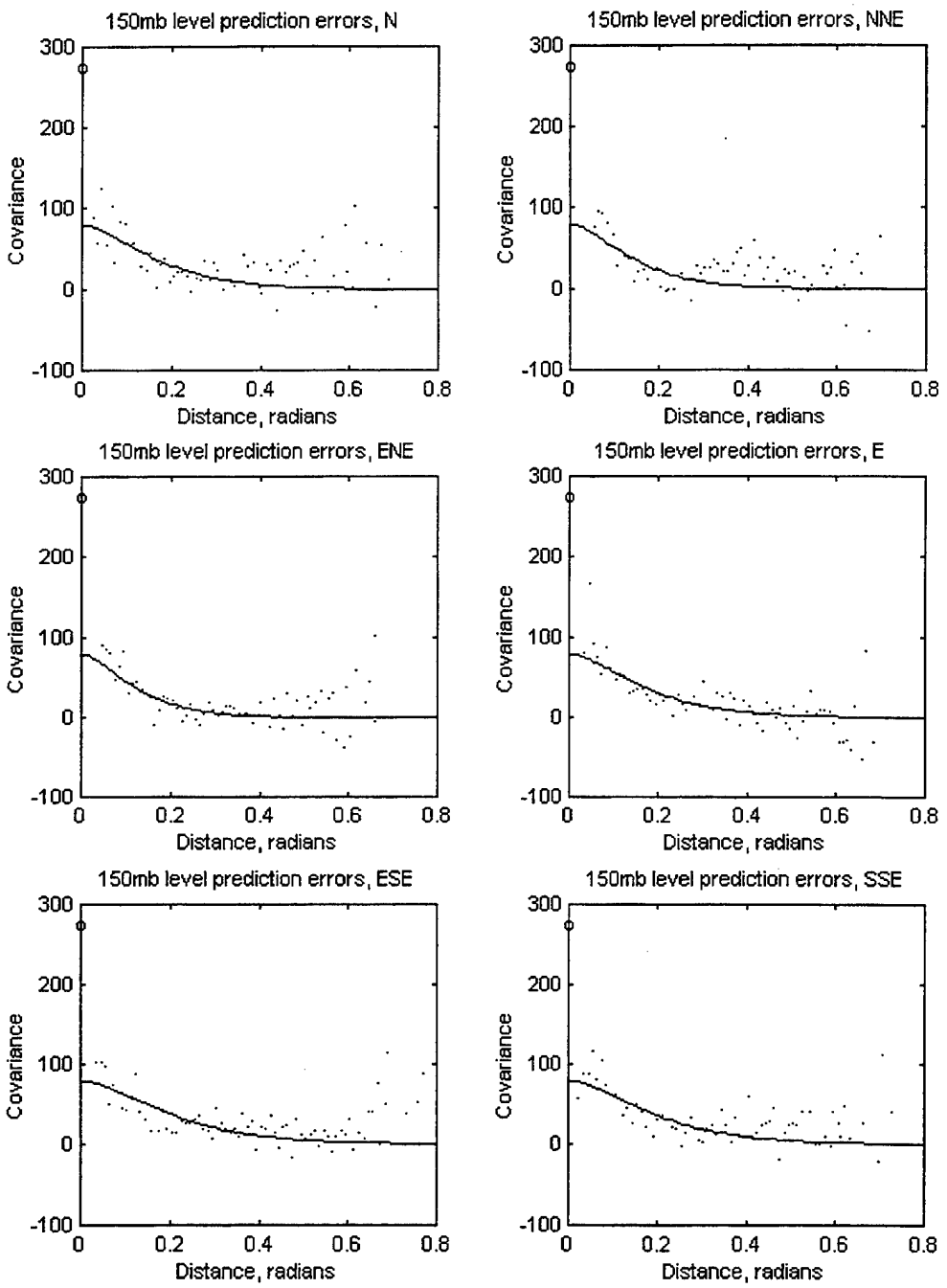

Figure 16: Binned pressure level error covariance data and spatial covariance approximations using SAR2 approximation, in six directions at 30° intervals, starting at North and going clockwise. This data is for the 150 mb level.

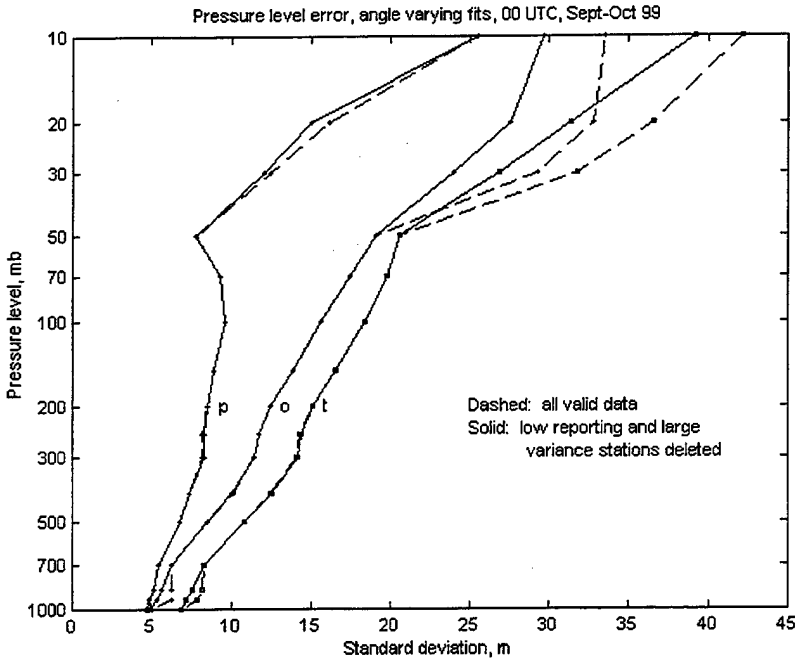

Figure 17: Standard deviations of prediction (p) and observation (o) errors as derived using the angle varying fit by SAR2 functions. The dashed line was derived using all valid covariance data, while the solid line used a data set that eliminated low reporting stations and large variance stations. The symbol t denotes the total standard deviation.

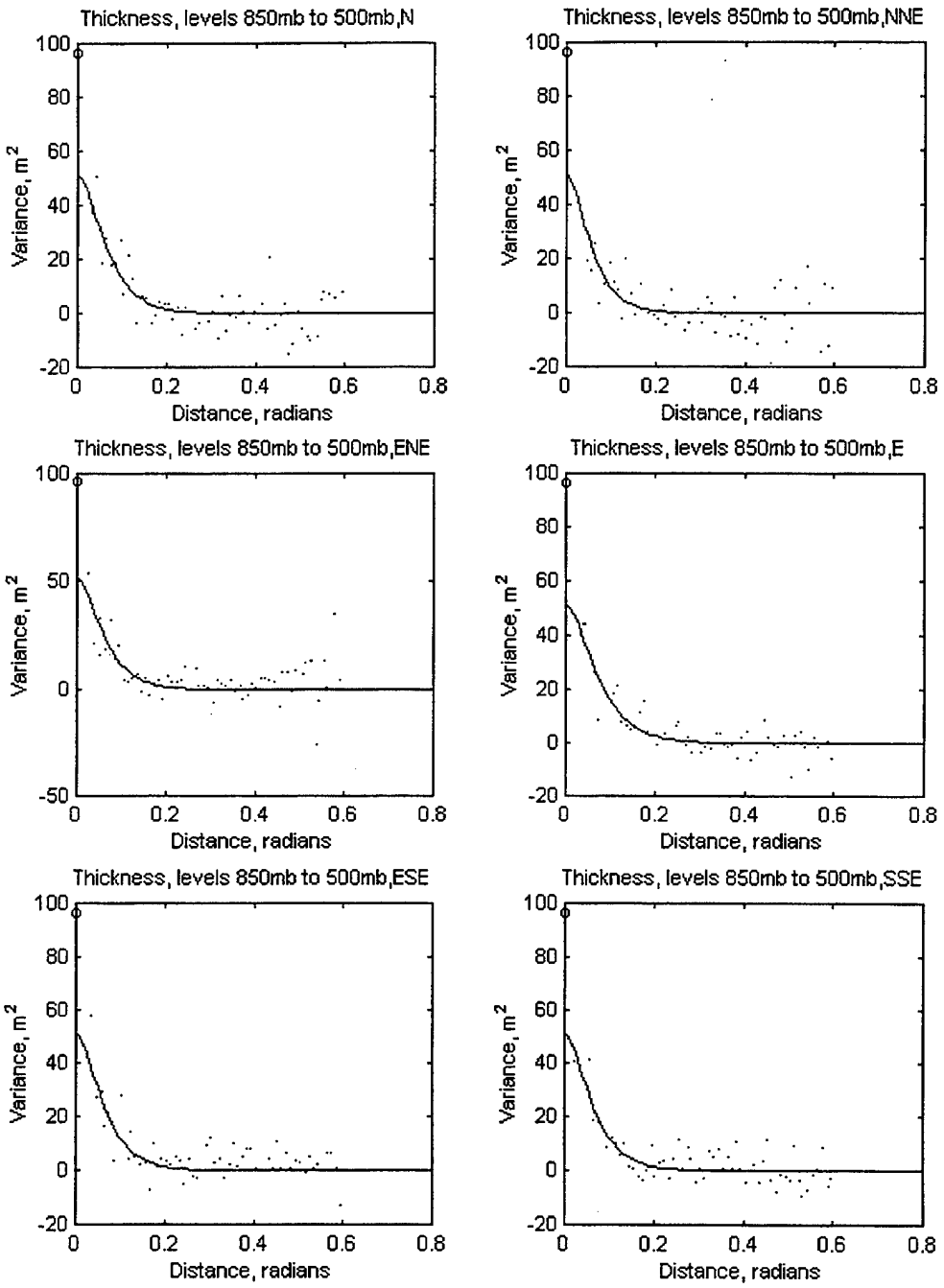

Figure 18: Pressure level thickness error covariance data and spatial covariance approximations fits using SAR2 approximation,starting at North and going clockwise. This data is for the 850 to 500 mb thickness.

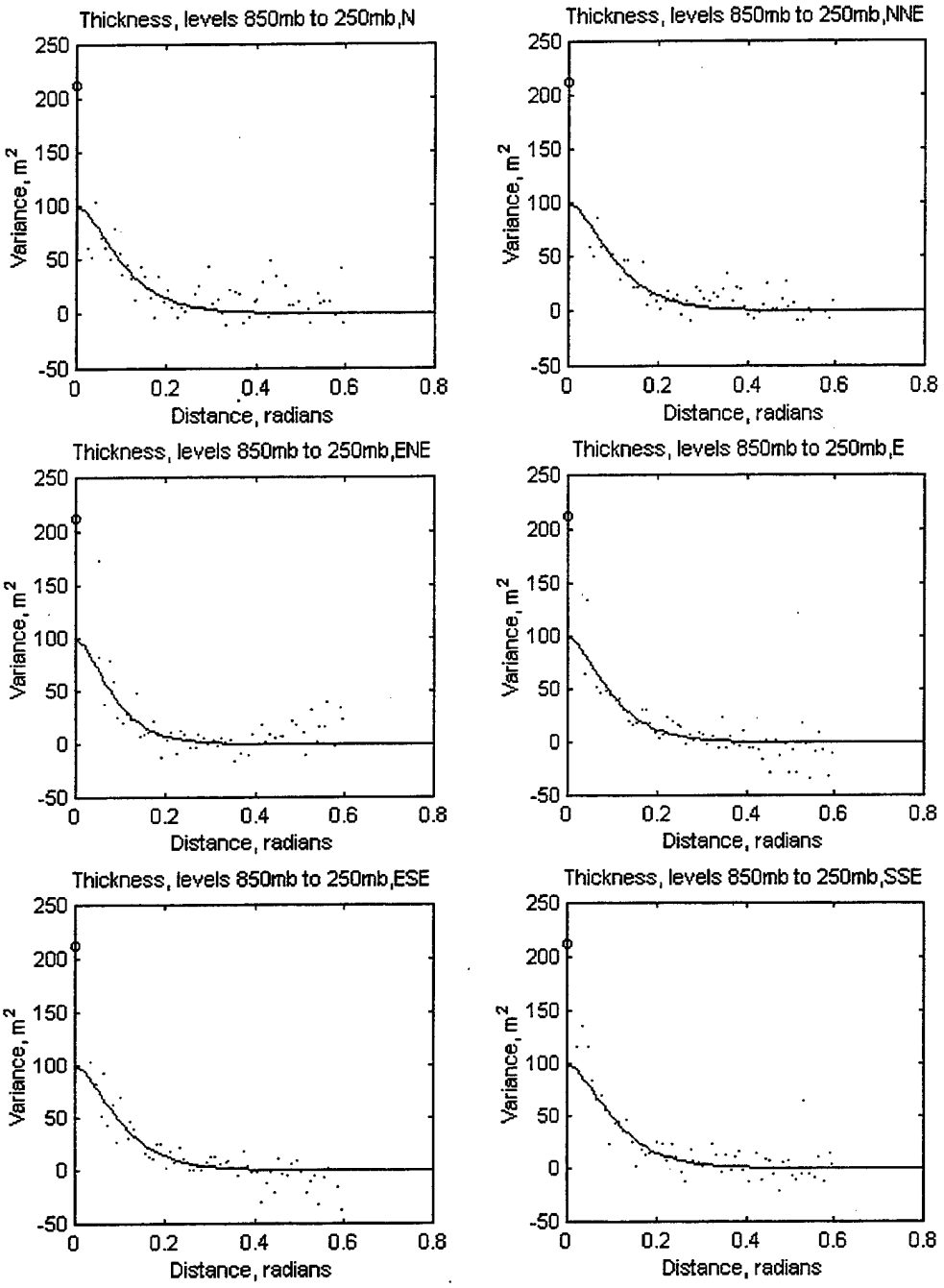

Figure 19: Pressure level thickness error covariance data and spatial covariance approximations fits using SAR2 approximation, in six directions at 30° intervals starting at North and going clockwise. This data is for the 850 to 250 mb thickness.

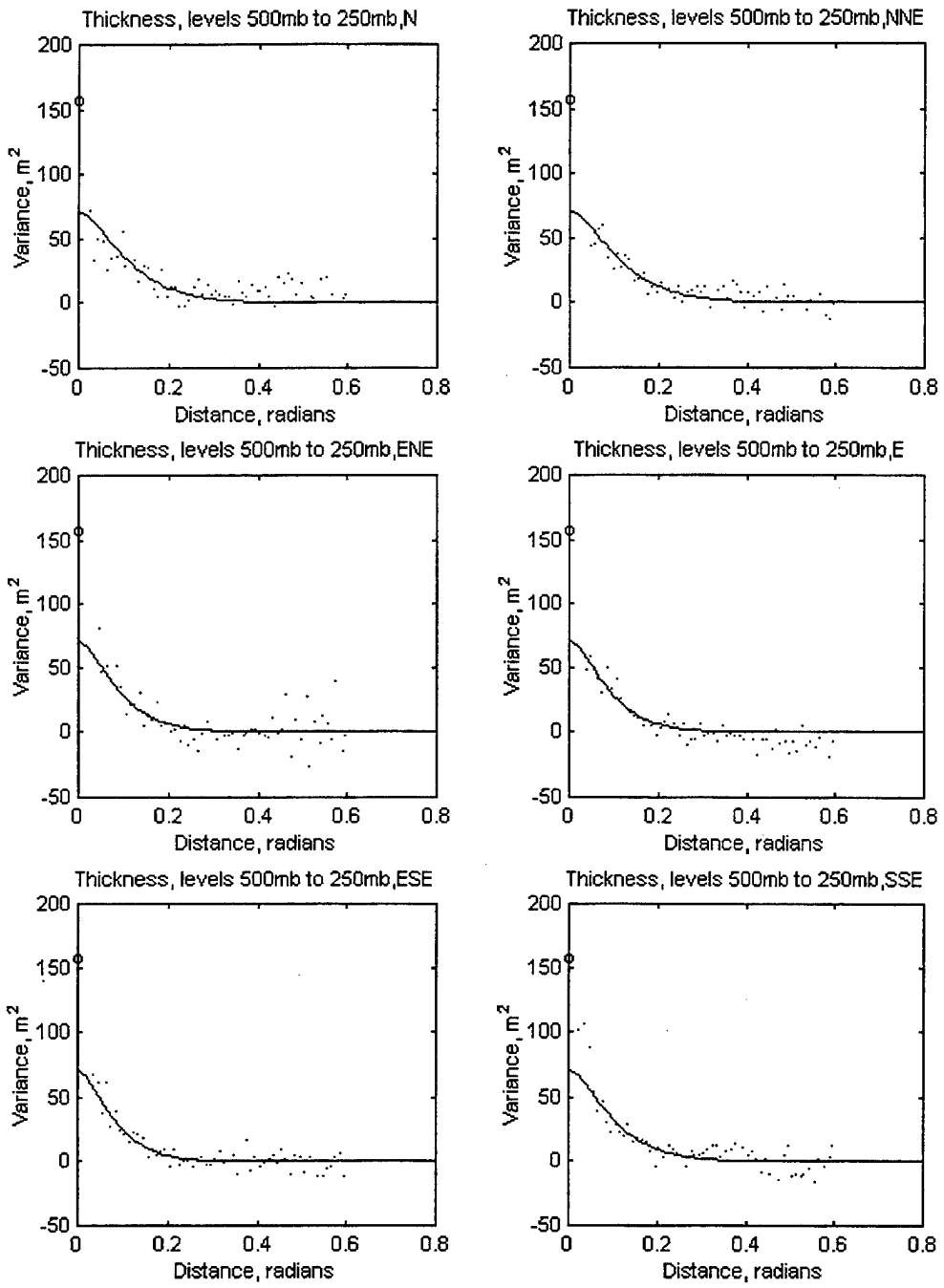

Figure 20: Pressure level thickness error covariance data and spatial covariance approximations fits using SAR2 approximation, in six directions at 30° intervals starting at North and going clockwise. This data is for the 500 to 250 mb thickness.

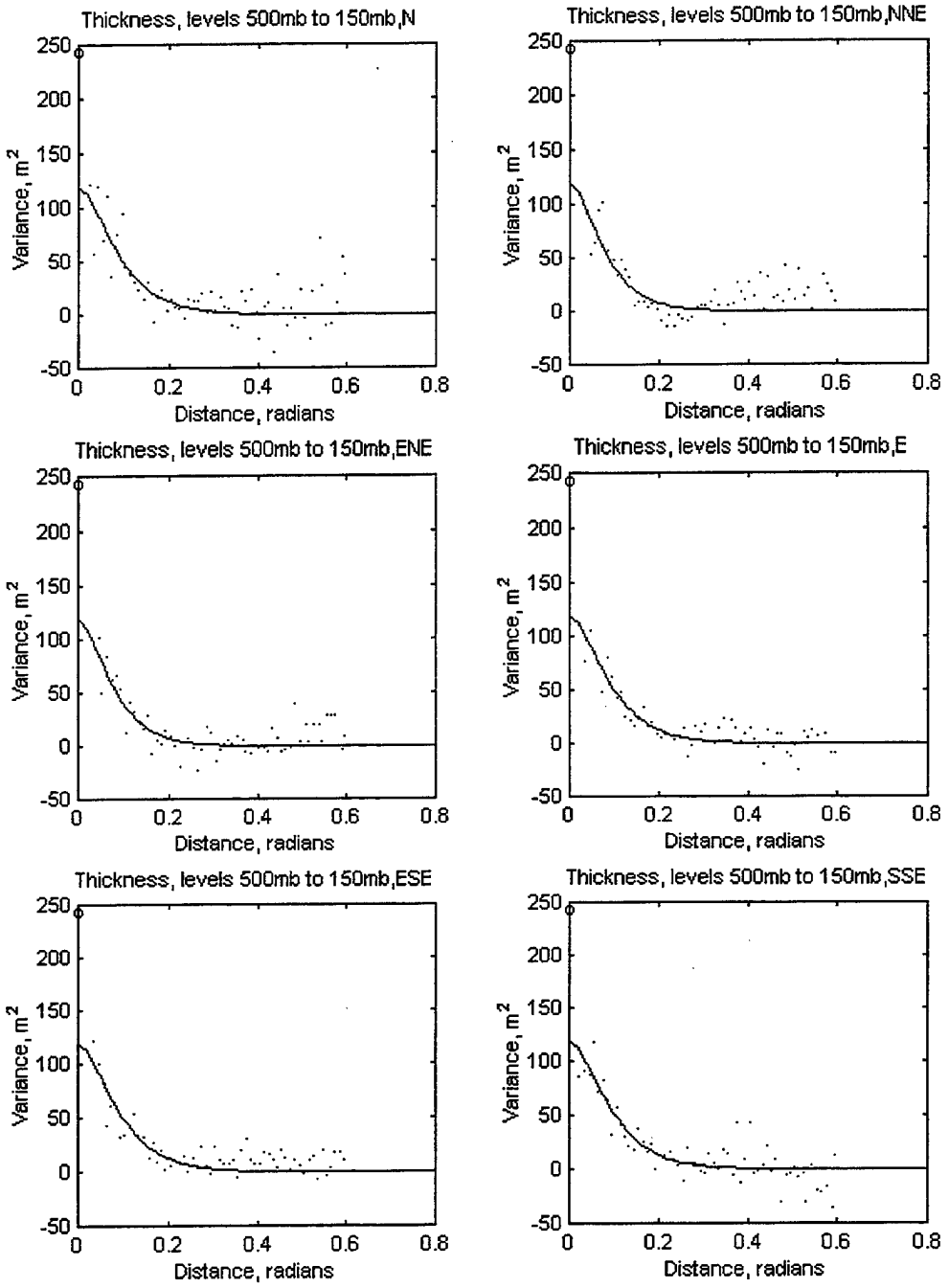

Figure 21: Pressure level thickness error covariance data and spatial covariance approximations using SAR2 function, in six directions at 30° intervals starting at North and going clockwise. The total innovation cross-covariance is shown by the "o" symbol on the axis. This data is for the 500 to 150 mb thickness.

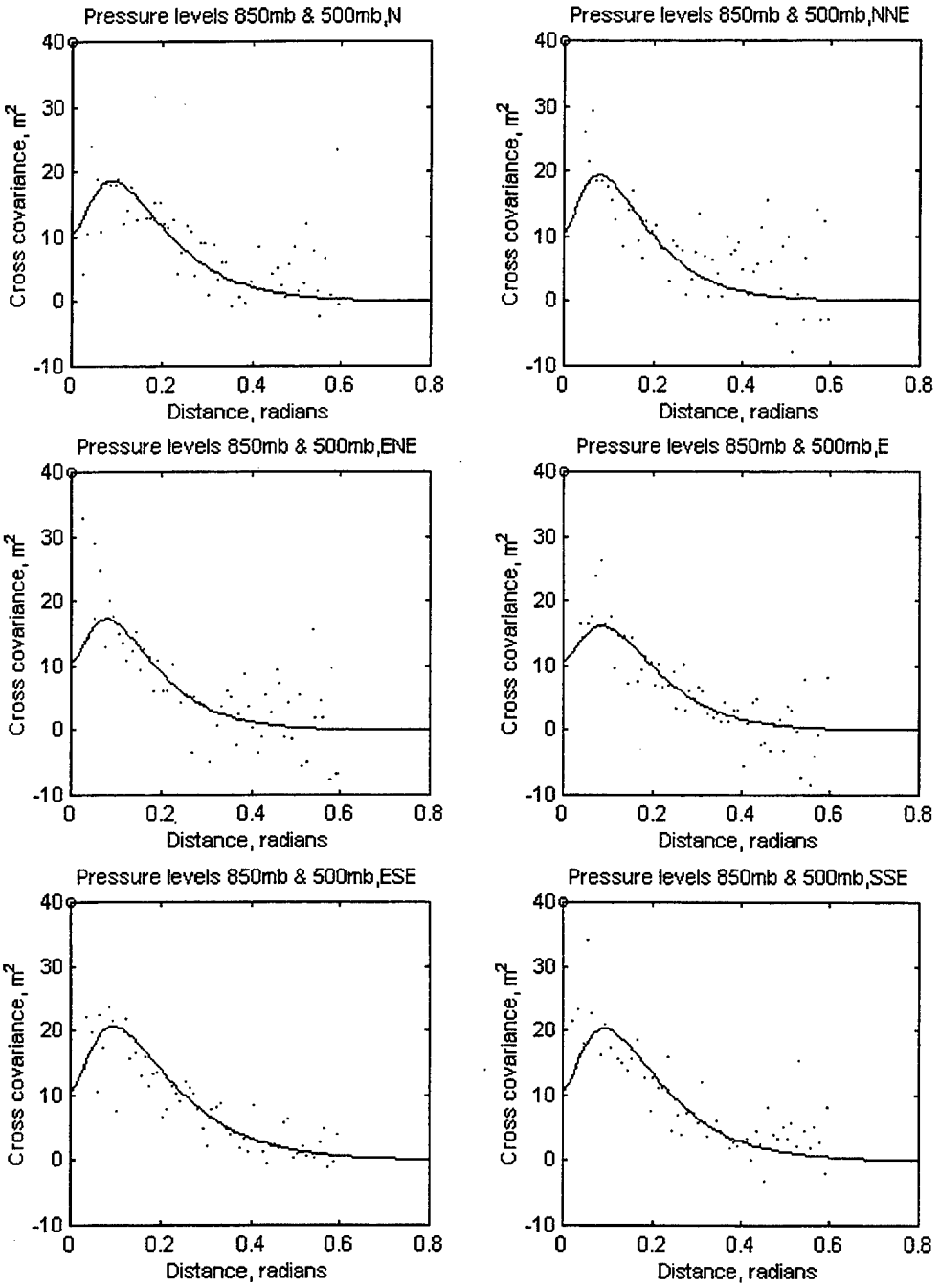

Figure 22: Interlevel pressure level error cross-covariance data from binned single level and binned interlevel thickness error data with implied fit from single level error fits and interlevel thickness error fits using SAR2 approximation. The total innovation cross-covariance is shown by the "o" symbol on the axis. This data is for the 850 and 500 mb levels.

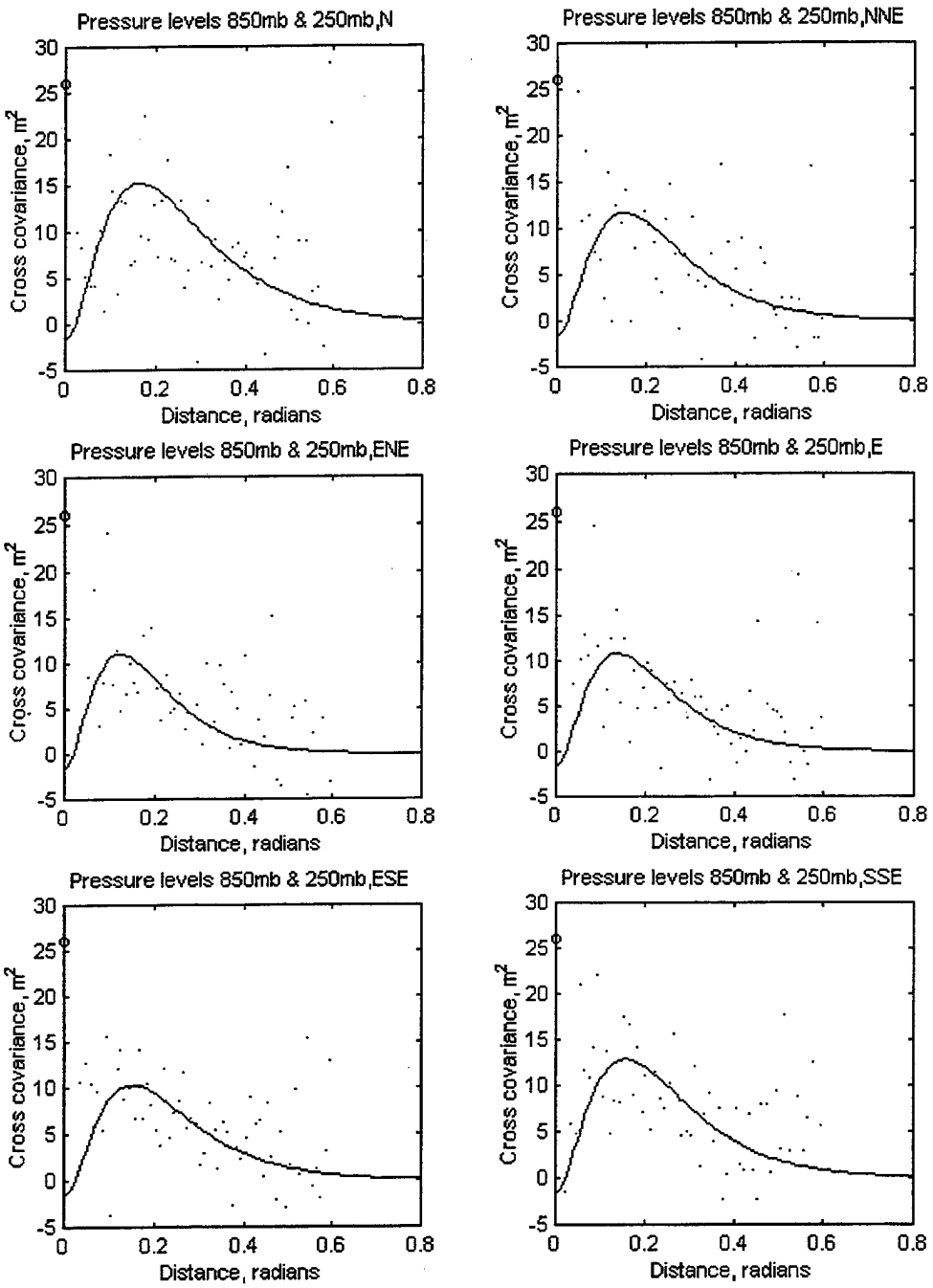

Figure 23: Interlevel pressure level error cross-covariance data from binned single level and binned interlevel thickness error data with implied fit from single level error fits and interlevel thickness error fits using SAR2 approximation. The total innovation cross-covariance is shown by the "o" symbol on the axis. This data is for the 850 and 250 mb levels.

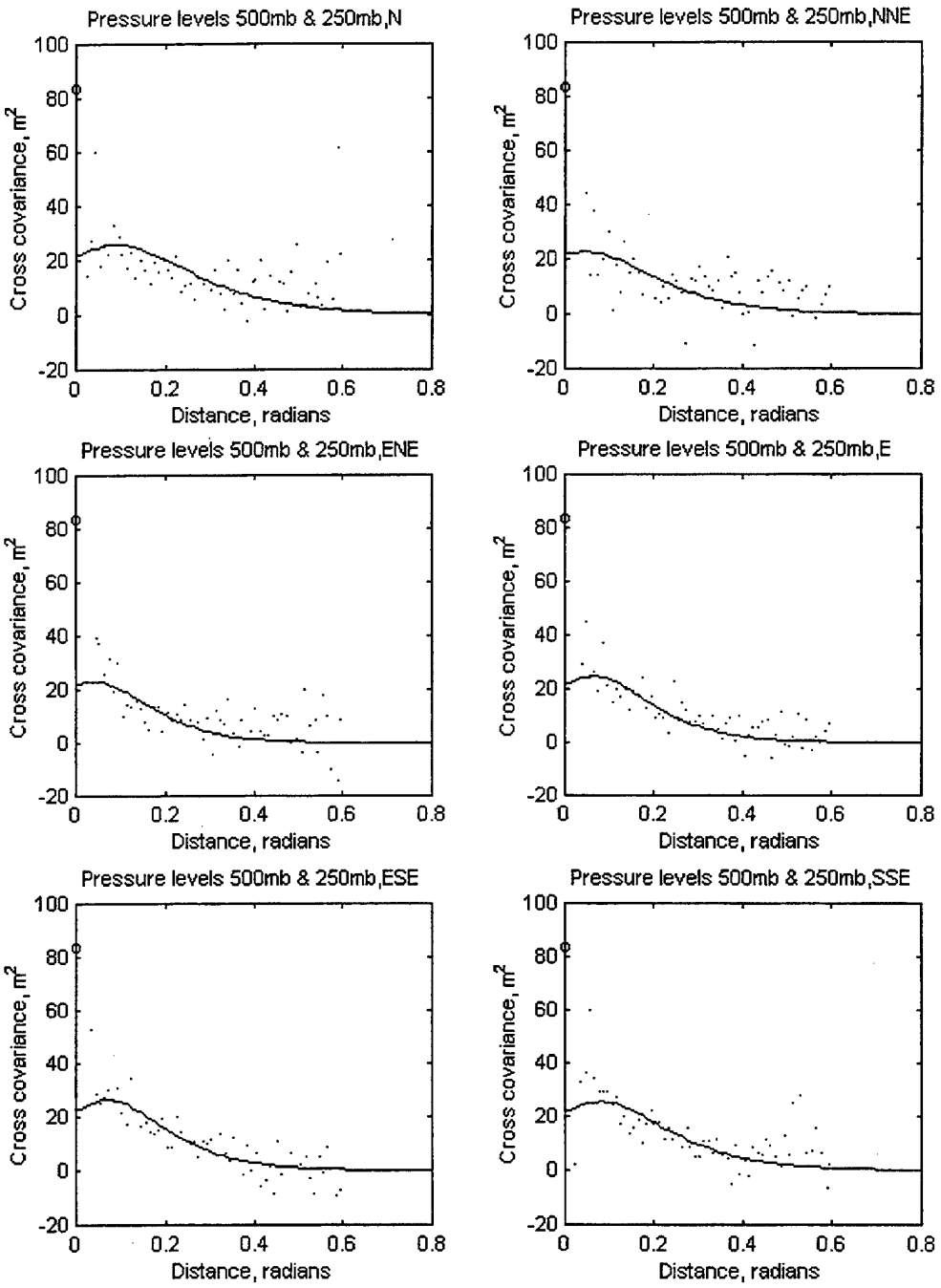

Figure 24: Interlevel pressure level error cross-covariance data from binned single level and binned interlevel thickness error data with implied fit from single level error fits and interlevel thickness error fits using SAR2 approximation. The total innovation cross-covariance is shown by the "o" symbol on the axis. This data is for the 500 and 250 mb levels.

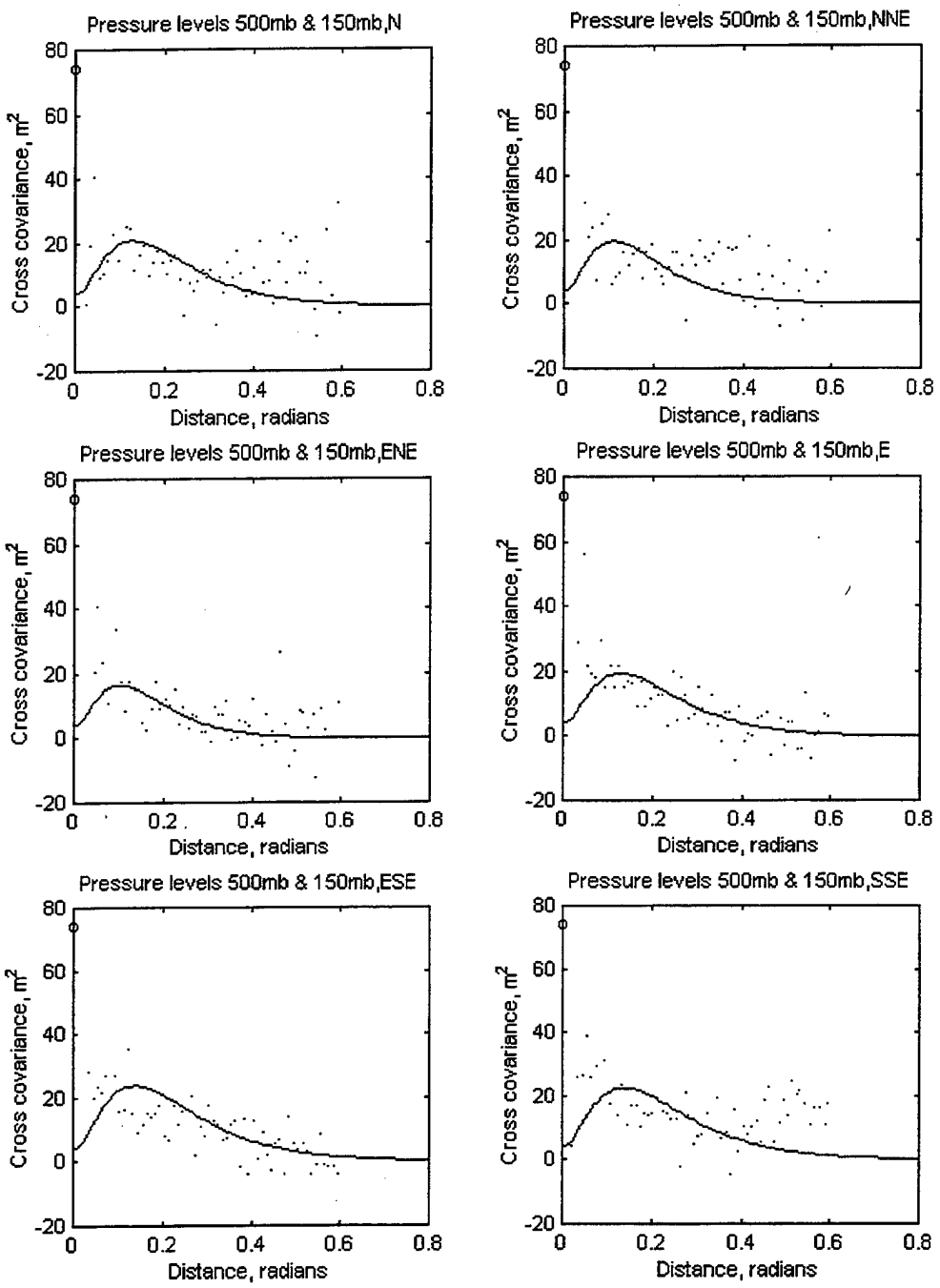

Figure 25: Interlevel pressure level error cross-covariance data from binned single level and binned interlevel thickness error data with implied fit from single level error fits and interlevel thickness error fits using SAR2 approximation. The total innovation cross-covariance is shown by the "o" symbol on the axis. This data is for the 500 and 150 mb levels.

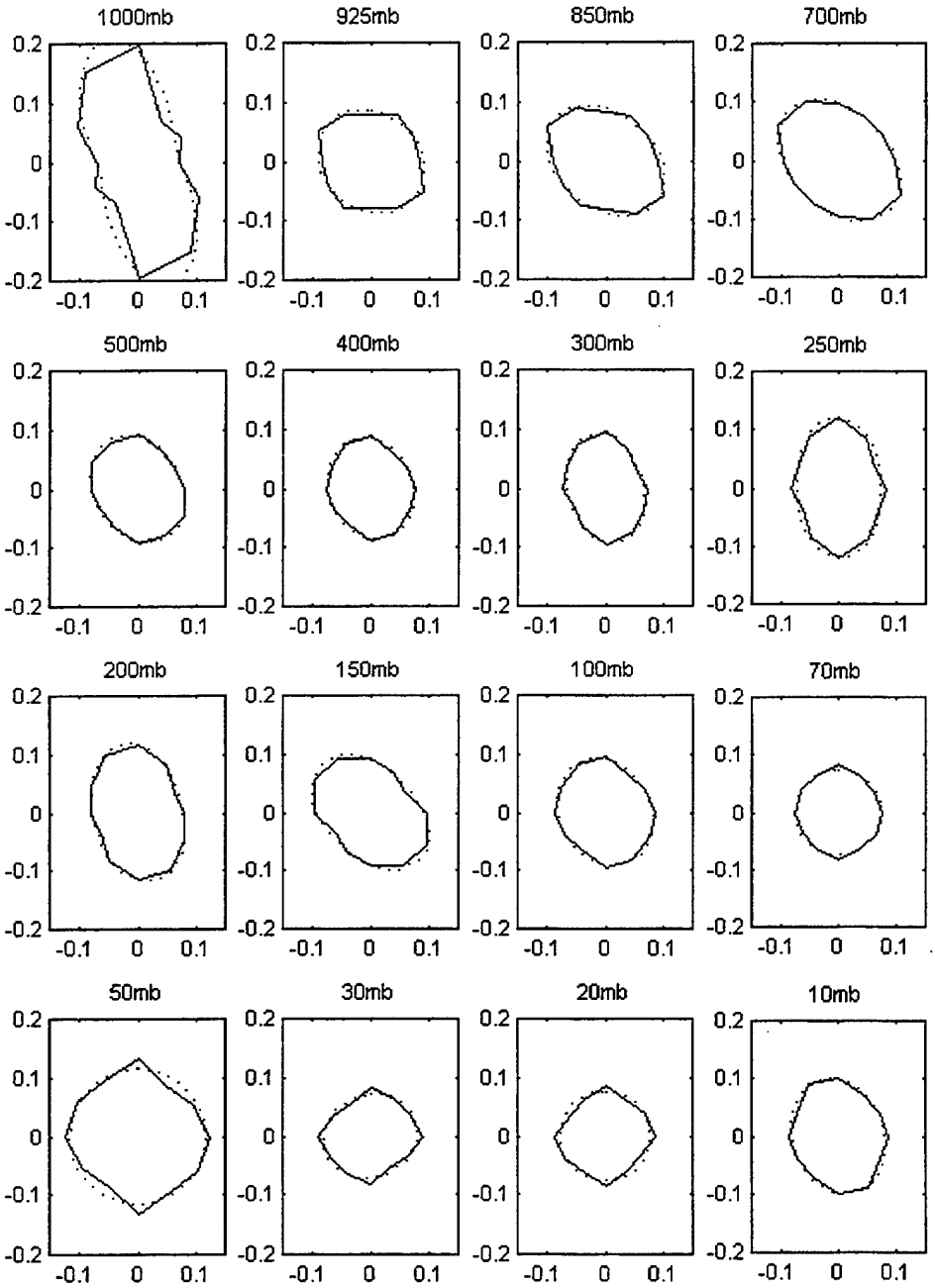

Figure 26: The solid lines give the correlation distances as a function of direction for the sixteen levels. The dashed lines represent the correlation distance as fit by a function of the form $D_l(\theta) = \dfrac{a_l}{(1+b_l \cos^2(\theta - \theta_l))}$ for level l.

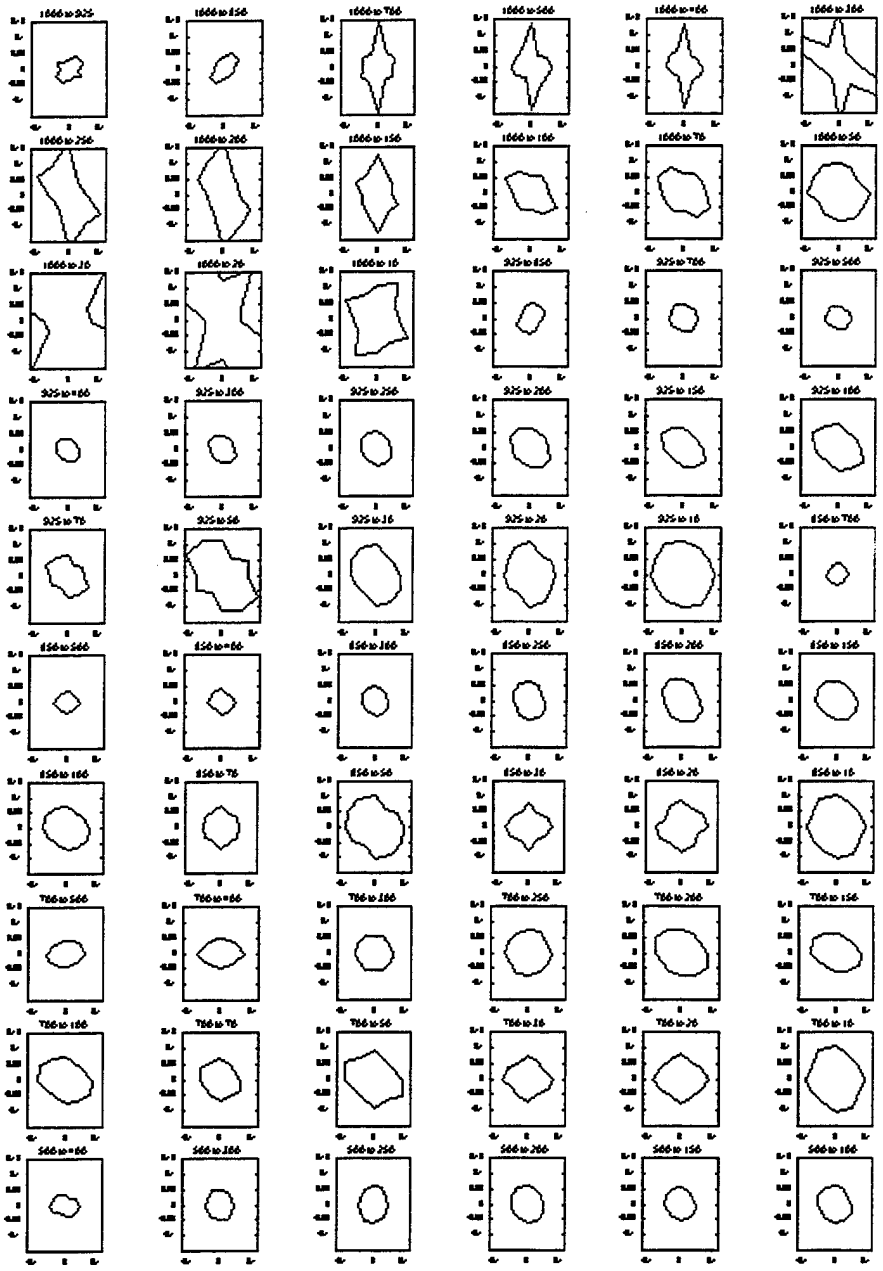

Figure 27a: Thumbnail plots of the correlation distance as a function of direction for interlevel thicknesses. Some distances are off-scale in order to have a reasonable common scale for all plots.

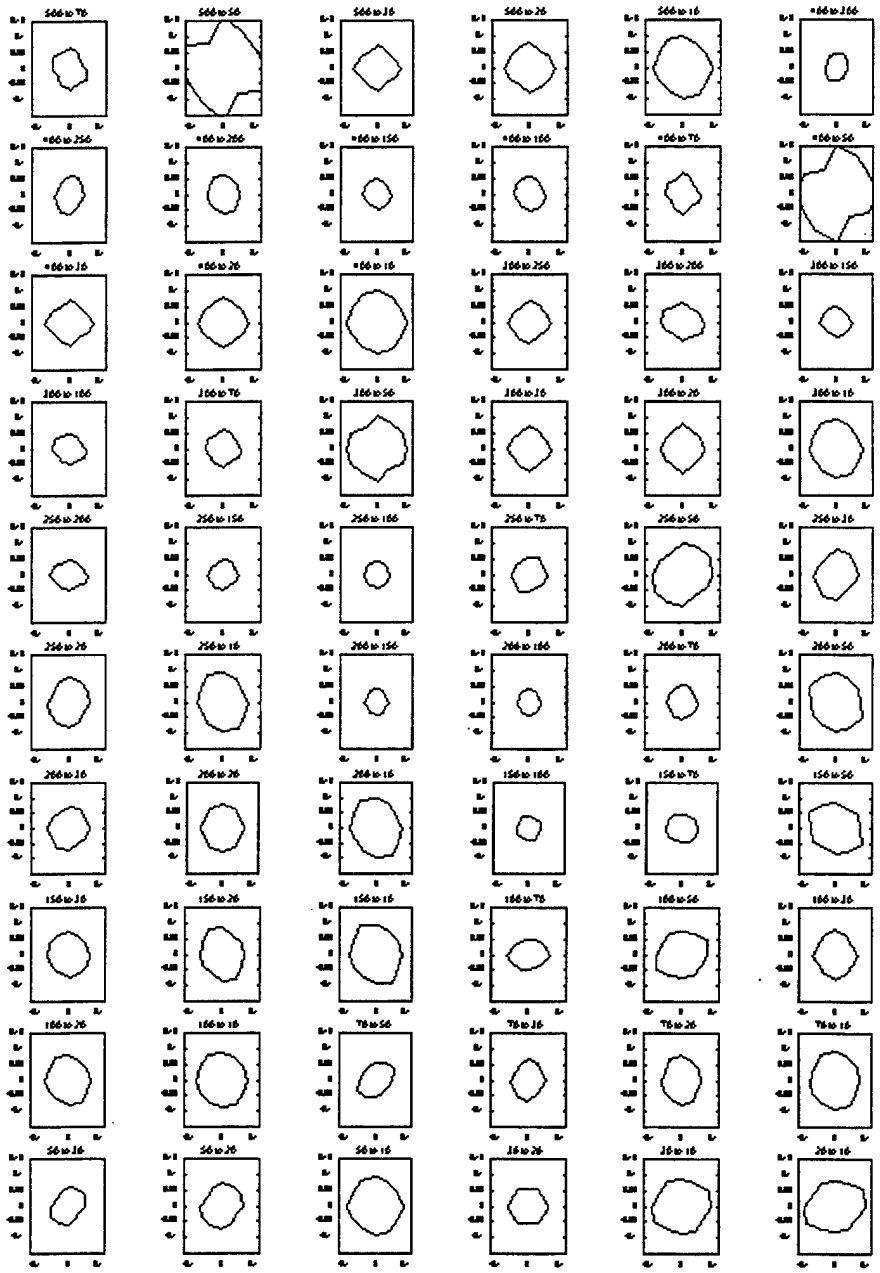

Figure 27b: Continuation of Figure 26a.

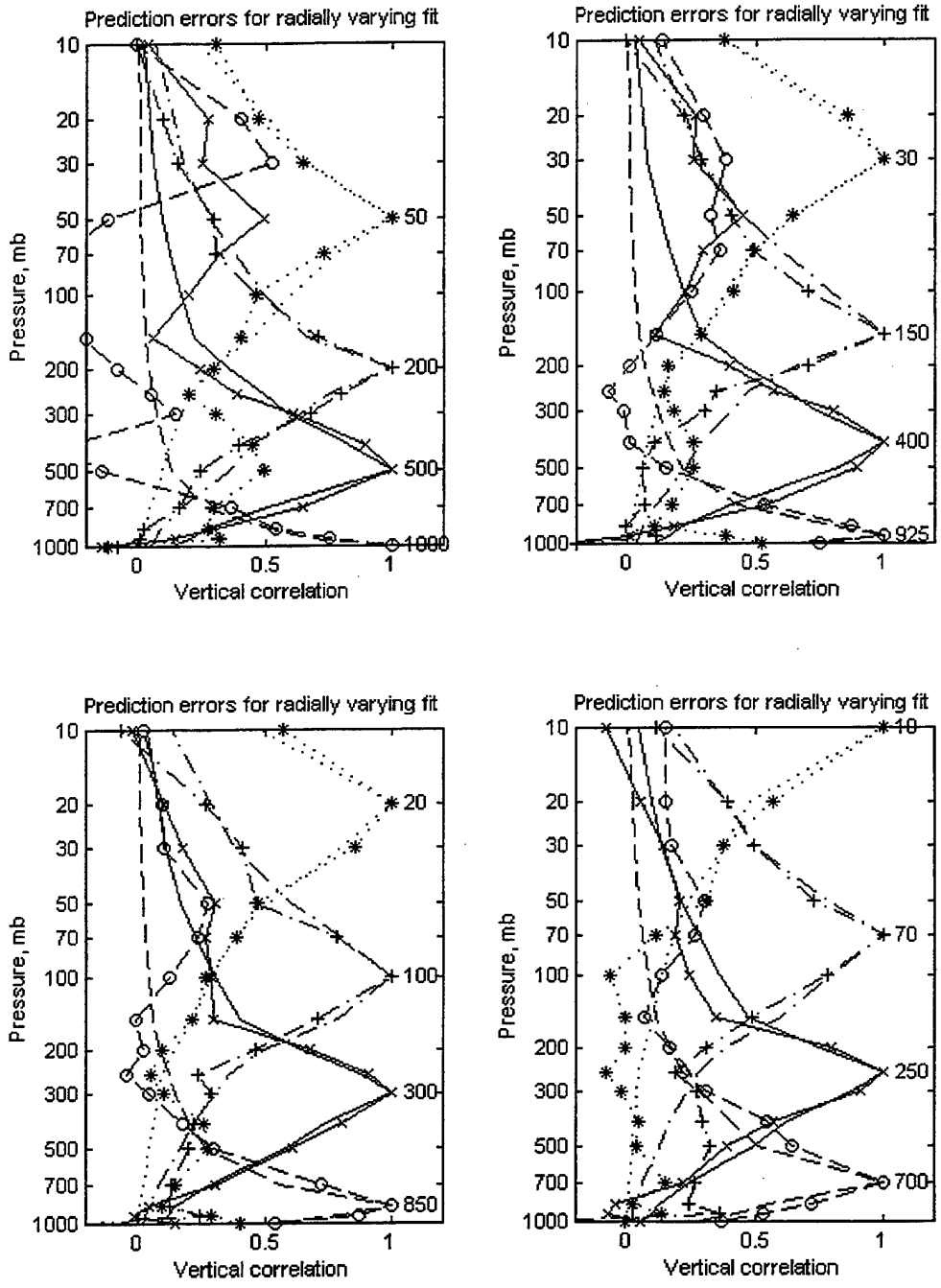

Figure 28: Vertical correlation curves for prediction error obtained from SAR2 fits (with symbols) and the approximation obtained by simultaneously finding a transformation of logP and fitting with a full third-order autoregressive (same type curve, without symbols).

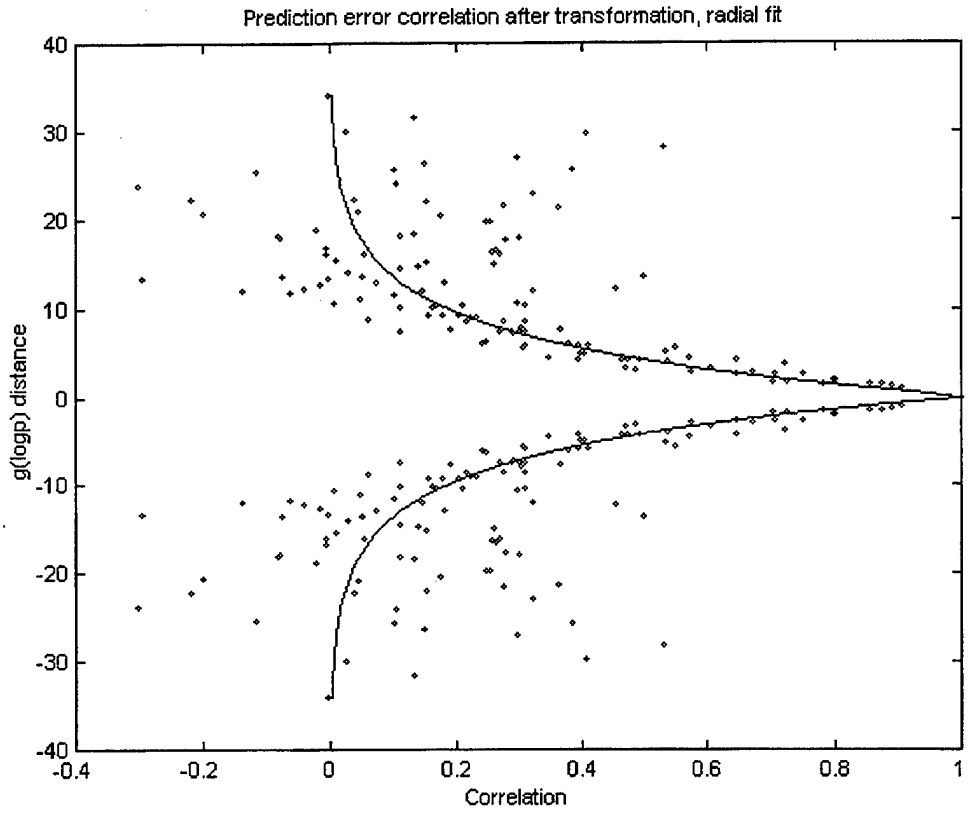

Figure 29: Vertical prediction error correlations and fitting curve for prediction error in the transformed coordinate system.

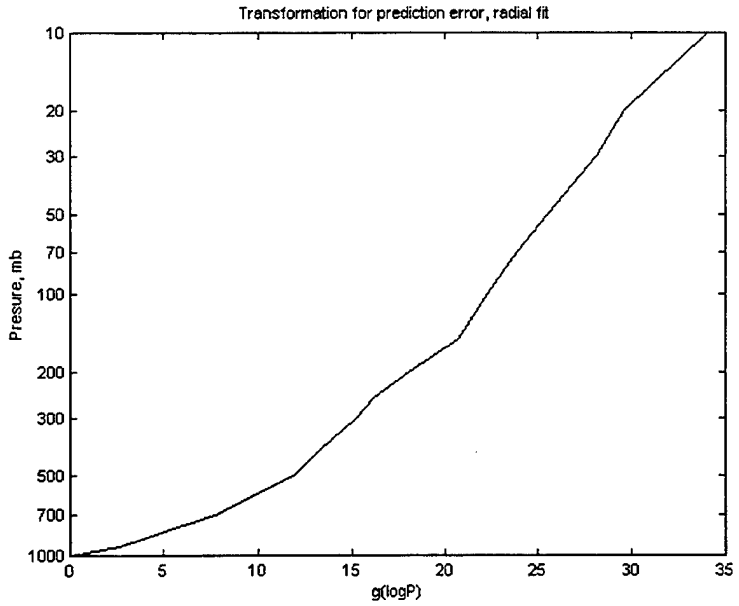

Figure 30: Transformation curve for prediction error.

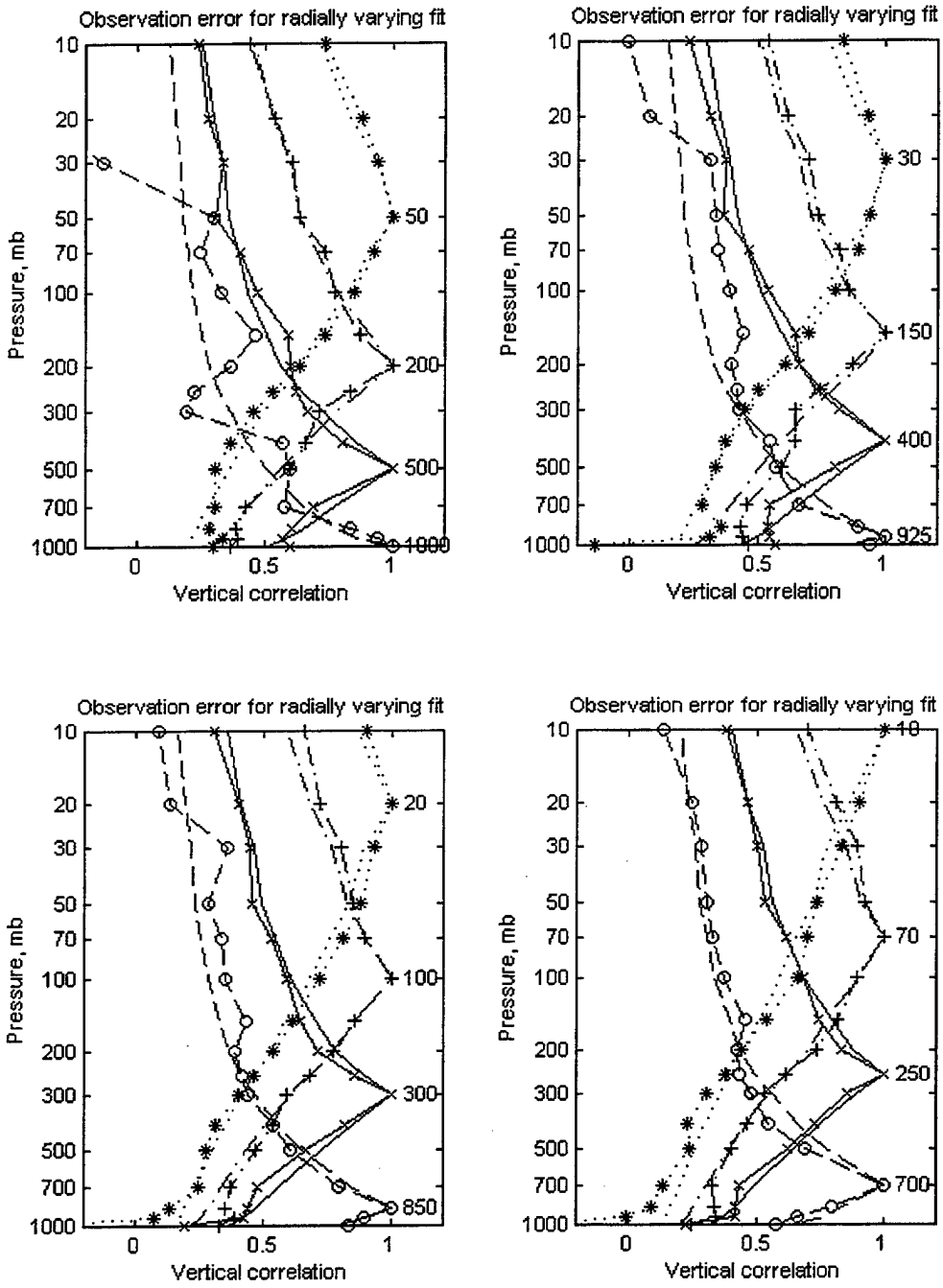

Figure 31: Vertical correlation curves for observation error obtained from SAR2 fits (with symbols) and the approximation obtained by simultaneously finding a transformation of logP and fitting with a full third-order autoregressive function (same type curve, without symbols).

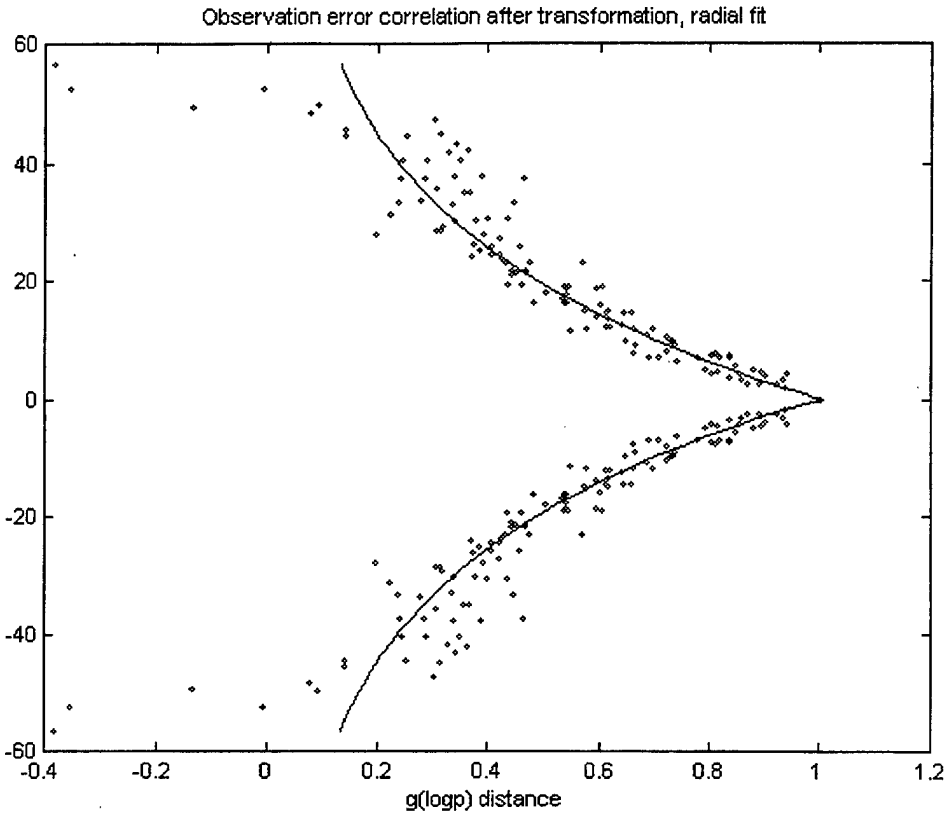

Figure 32: Vertical observation error correlations and fitting curve for observation error in the transformed coordinate system.

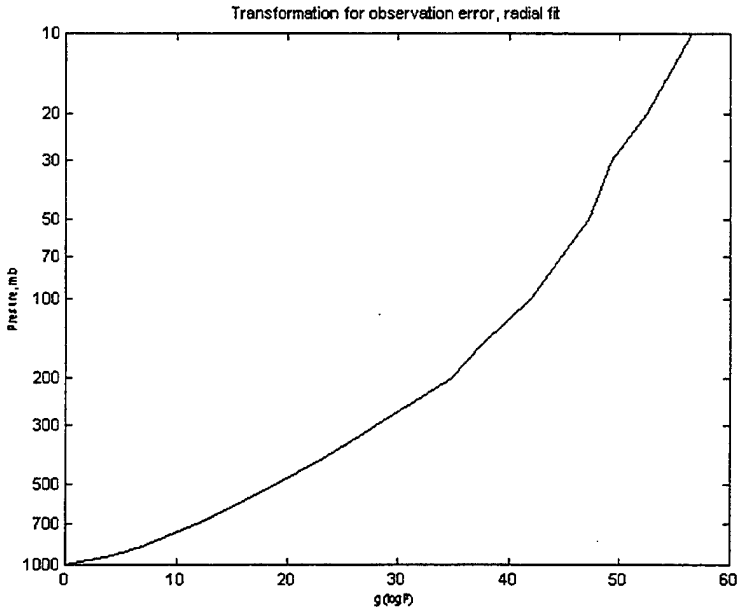

Figure 33: Transformation curve for observation error.

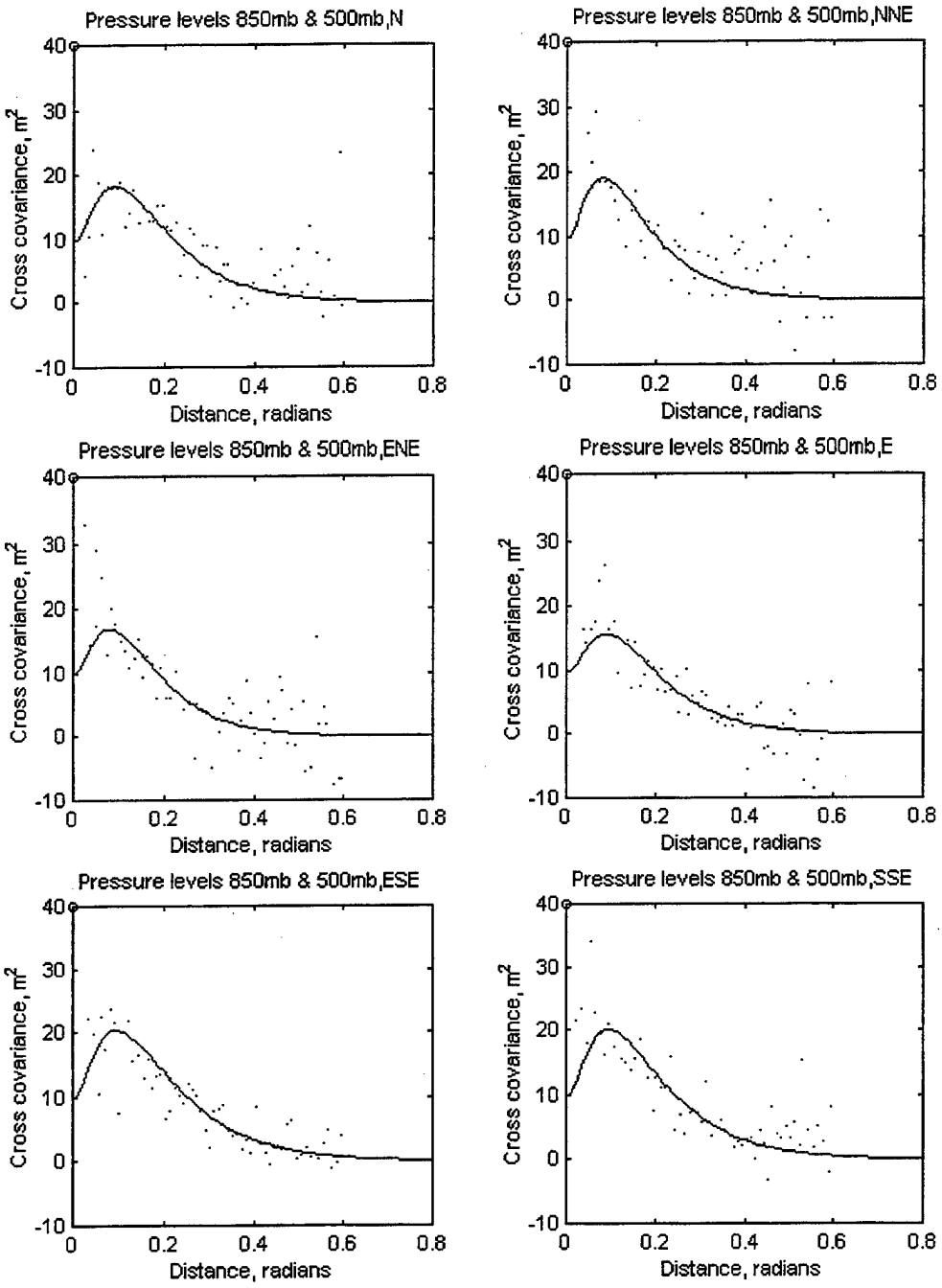

Figure 34: Interlevel pressure level error cross-covariance data from binned single level and binned interlevel thickness error data with implied fit from single level error fits and interlevel thickness error fits using SAR2 polynomials after imposing the implied thickness variance derived from the vertical correlation approximations shown in Figure 28 . This data is for the 850 and 500 mb levels. Compare with Figure 22.

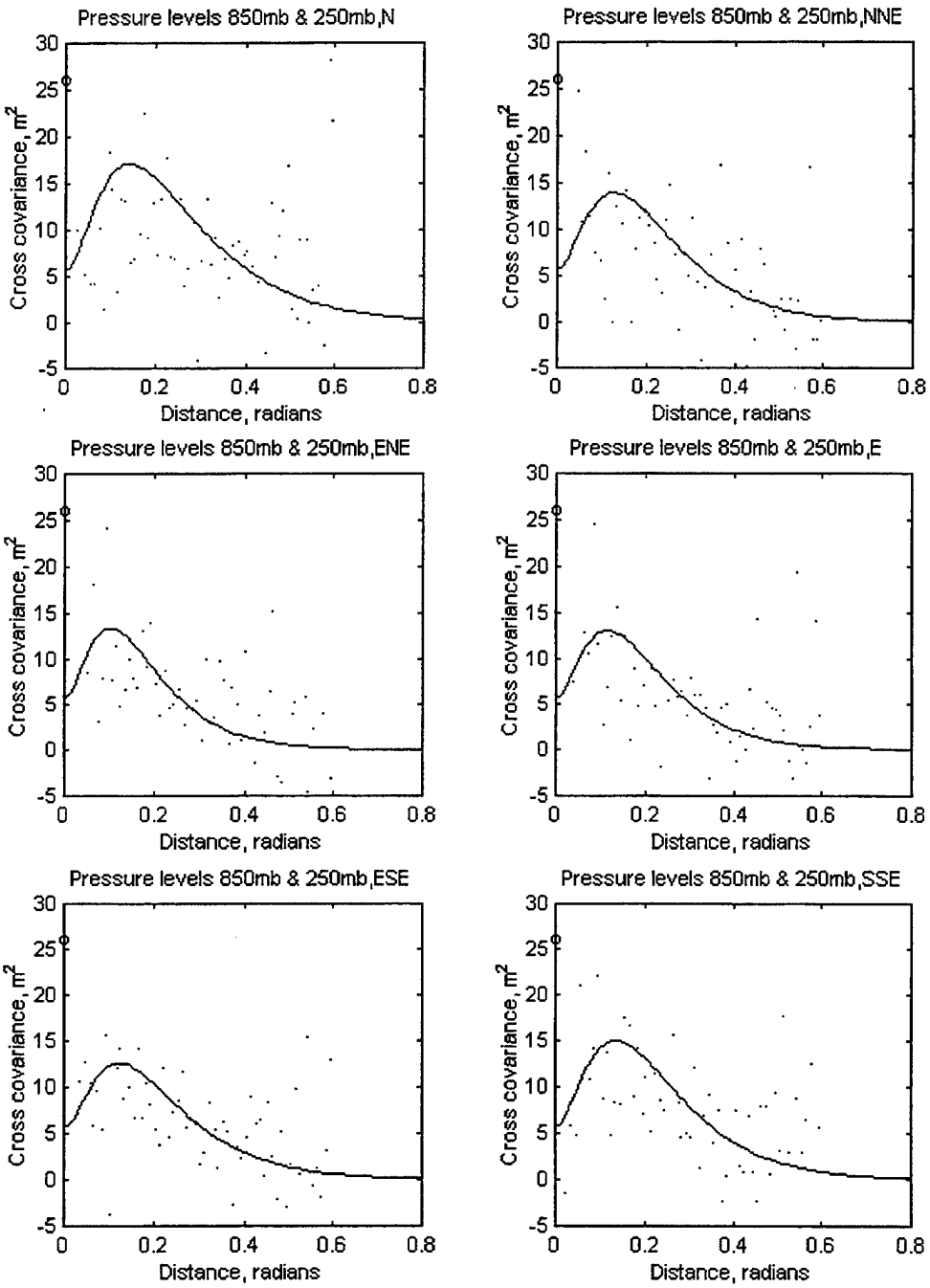

Figure 35: Interlevel pressure level error cross-covariance data from binned single level and binned interlevel thickness error data with implied fit from single level error fits and interlevel thickness error fits using SAR2 polynomials after imposing the implied thickness variance derived from the vertical correlation approximations shown in Figure 28 . This data is for the 850 and 250 mb levels. Compare with Figure 23.

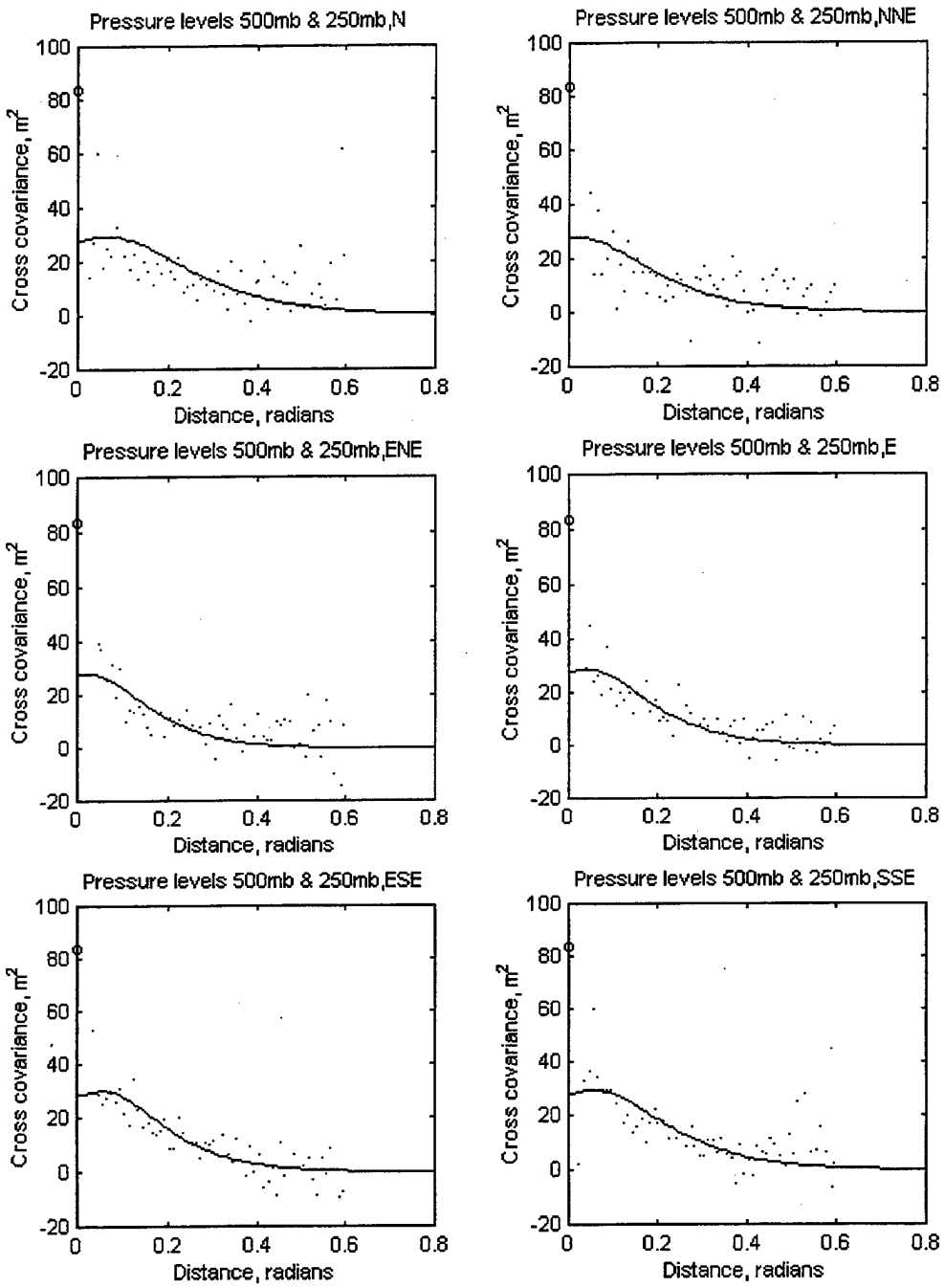

Figure 36: Interlevel pressure level error cross-covariance data from binned single level and binned interlevel thickness error data with implied fit from single level error fits and interlevel thickness error fits using SAR2 polynomials after imposing the implied thickness variance derived from the vertical correlation approximations shown in Figure 28 . This data is for the 500 and 250 mb levels. Compare with Figure 24.

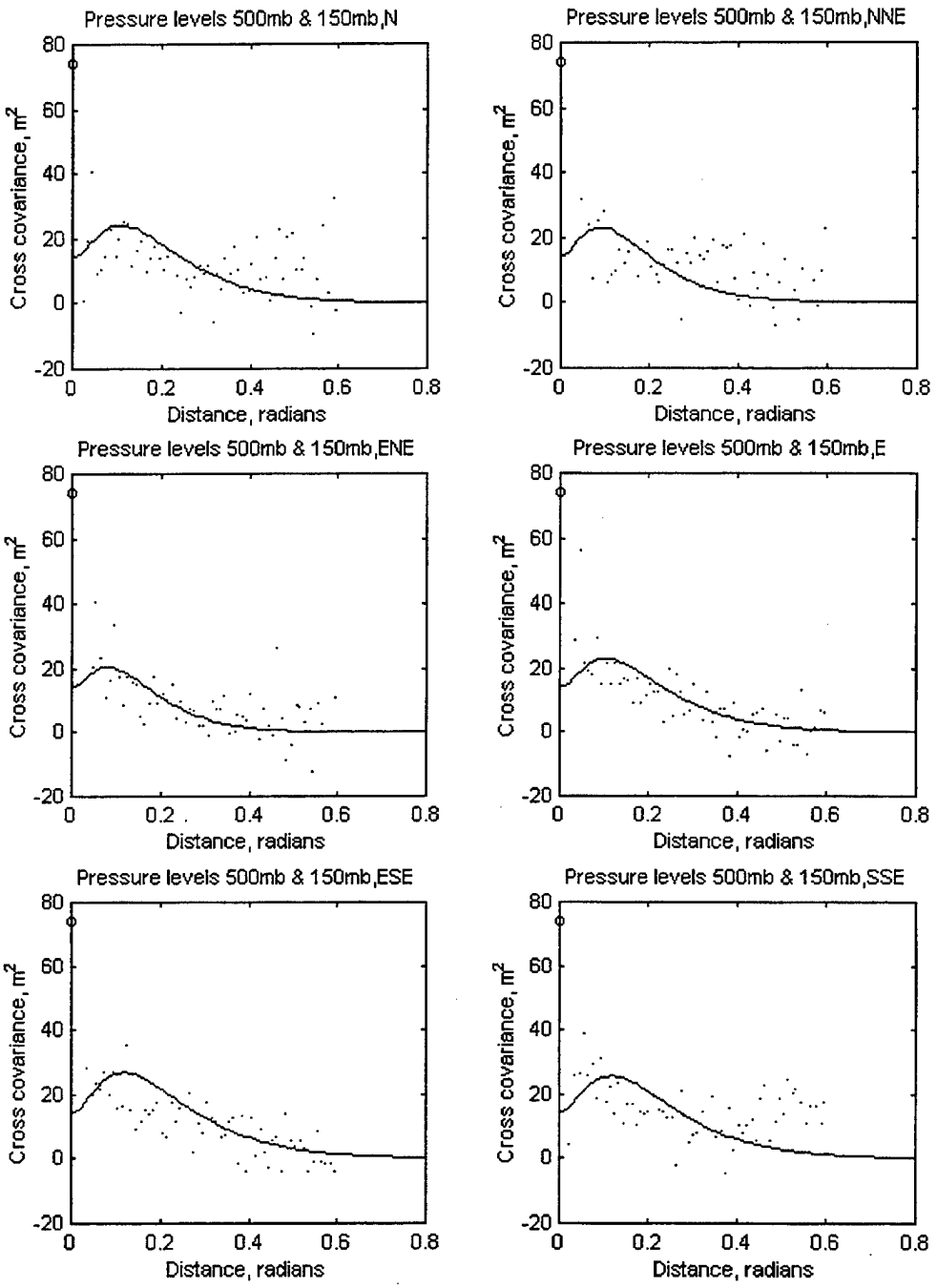

Figure 37: Interlevel pressure level error cross-covariance data from binned single level and binned interlevel thickness error data with implied fit from single level error fits and interlevel thickness error fits using SAR2 polynomials after imposing the implied thickness variance derived from the vertical correlation approximations shown in Figure 28 . This data is for the 500 and 150 mb levels. Compare with Figure 25.